이학박사 학위논문

Development and Analytical Validation of Clinical Assay for AFP, AFP-L3 and PIVKA-II using Multiple Reaction Monitoring-Mass Spectrometry (MRM-MS)

다중반응검지법을 이용한 AFP, AFP−L3, PIVKA−II 의 임상 정량법 개발과 분석적 검증

2019년 08월

서울대학교 대학원

의과학과 의과학전공

손 아 름

A thesis of the Degree of Doctor of Philosophy

다중반응검지법을 이용한 AFP, AFP-L3, PIVKA-II 의 임상 정량법 개발과 분석적 검증

Development and Analytical Validation of Clinical Assay for AFP, AFP-L3 and PIVKA-II using Multiple Reaction Monitoring-Mass Spectrometry (MRM-MS)

August 2019

Major in Biomedical Sciences

Department of Biomedical Sciences

Seoul National University

Graduate School

Areum Sohn

다중반응검지법을 이용한
AFP, AFP-L3, PIVKA-II 의
임상 정량법 개발과 분석적 검증

지도교수 김 영 수

이 논문을 이학박사 학위논문으로 제출함

2019년 04월

서울대학교 대학원

의과학과 의과학전공

손 아 름

손아름의 이학박사 학위논문을 인준함

2019년 07월

위 원 장 박 태 성 (인)

부위원장 김 영 수 (인)

위 원 이 유 진 (인)

위 원 권 용 태 (인)

위 원 김 경 곤 (인)

Development and Analytical Validation of Clinical Assay for AFP, AFP-L3 and PIVKA-II using Multiple Reaction Monitoring-Mass Spectrometry (MRM-MS)

by

Areum Sohn

A thesis submitted to the Department of Biomedical Science in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Biomedical Sciences at Seoul National University College of Medicine

July 2019

Approved by Thesis Committee:

Professor _____ Chairman

Professor Youngsoo Kim Vice chairman

Professor Eugene Yi

Professor Yong Tae Kwon

Professor Kyungsoo Kim

ABSTRACT

Development and Analytical Validation of Clinical Assay for AFP, AFP-L3 and PIVKA-II using Multiple Reaction Monitoring-Mass Spectrometry (MRM-MS)

Areum Sohn

Major in Biomedical Sciences

Department of Biomedical Sciences

Seoul National University

Graduate School

Introduction: Alpha-fetoprotein (AFP), lens culinaris agglutinin-reactive fraction of alpha-fetoprotein (AFP-L3) protein induced by vitamin K absence or antagonist-II (PIVKA-II) are serum biomarkers for hepatocellular carcinoma (HCC). These markers have been measured typically by liquid-phase binding assay (LiBA). However, LiBA does not always reflect accurate concentration, due to its low analytical sensitivity. Thus, I aimed to develop an analytically sensitive multiple reaction monitoring-mass spectrometry (MRM-MS) assay to quantify AFP, AFP-L3 and PIVKA-II in serum.

Methods: The assay entailed the addition of a stable isotope-labeled internal standard protein analog into the serum, the enrichment of AFP, and prothrombin using a monoclonal antibody, the fractionation of AFP-L3 using lens culinaris agglutinin lectin, deglycosylation for AFP-L3 fractionation, trypsin digestion (AFP, AFP-L3), chymotrypsin (PIVKA-II), online desalting, and MRM-MS analysis. The performance of the MRM-MS assay was compared with that of LiBA in 400 human serum samples (100 chronic hepatitis, 100 liver cirrhosis, and 200 HCC group). Subsequently, the assays for AFP, AFP-L3 and PIVKA-II were analytically validated per the US Food and Drug Administration (FDA), European medicines agency (EMA), Korea FDA (KFDA), and Clinical & Laboratory Standards Institute (CLSI) for clinical implementation.

Results: The lower limit of quantification (LLOQ) of the MRM-MS assay (0.051 ng/mL) for AFP, AFP-L3 was below that of LiBA. Thus, AFP-L3 values, which measured by LiBA in HCC samples (n=39), were detected by the MRM-MS assay. For PIVKA-II, the representative signature peptide was selected to measure PIVKA-II concentrations by MRM-MS. The linearity ranged from 1.28 ng/mL to 100000 ng/mL. In additions, the MRM-MS assay for AFP, AFP-L3 and PIVKA-II was validated to meet almost criteria for 12 categories (calibration curve, analytical specificity (selectivity or interference), analytical sensitivity, carryover, precision, recovery of assay, matrix effect, recovery of immunoprecipitation, dilution integrity, stability, reproducibility, and quality control (QC) of samples and frequency) to

confirm whether the assay meet the international guidelines.

Conclusions: I developed methods for quantifying AFP, AFP-L3 in serum by MRM-MS-based assay that can overcome the low analytical sensitivity of LiBA. For PIVKA-II, I identified the surrogate peptide for PIVKA-II levels in human serum, and established MRM-MS assay for the first time. For the confirmation of robustness and reproducibility, the developed MRM-MS assays for AFP, AFP-L3 and PIVKA-II were validated per the US Food and Drug Administration (FDA), European Medicines Agency (EMA), Korea FDA (KFDA), and Clinical and Laboratory Standards Institute (CLSI).

Keywords: Multiple Reaction Monitoring; Biomarker; Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein (AFP-L3); Protein induced by vitamin K absence or antagonist-II (PIVKA-II); Method validation; Hepatocellular Carcinoma; Alpha-Fetoprotein

Student number: 2014-21997

* This work is published in Analytical and Bioanalytical Chemistry Journal. A quantitative analytical method for PIVKA-II using multiple reaction monitoring-mass spectrometry for early diagnosis of hepatocellular carcinoma (A. sohn, H. Kim,

S.J. Yu, J.H. Yoon and Y. Kim). Published 6 Feb 2017/ Analytical and Bioanalytical Chemistry 10.1007/s00216-017-0227-8.

* This work is published in Journal of Pharmaceutical and Biomedical Analysis Journal. Fully validated SRM-MS-based method for absolute quantification of PIVKA-II in human serum: Clinical applications for patients with HCC (A. sohn, H. Kim, I. Yeo, Y. Kim, M. Son, S.J. Yu, J.H. Yoon and Y. Kim). Published 21 Apr 2018/ Journal of Pharmaceutical and Biomedical Analysis Journal 10.1016/j.jpba.2018.04.025.

* This work is published in Clinical Chemistry. Clinical Assay for AFP-L3 by Using Multiple Reaction Monitoring–Mass Spectrometry for Diagnosing Hepatocellular Carcinoma (H. Kim, A. sohn, I. Yeo, S.J. Yu, J.H. Yoon and Y. Kim). Published 27 Jul 2018/ Clinical chemistry 10.1373/clinchem.2018.289702.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFP: α-fetoprotein

HCC: hepatocellular carcinoma

LCA: Lens culinaris agglutinin

AFP-L3: Lens culinaris agglutinin-reactive fraction of α-fetoprotein

PIVKA-II: Prothrombin induced by vitamin K absence-II

LiBA: liquid-phase binding assay

μTAS: micro-total assay system

FDA: Food and Drug Administration

MRM-MS: multiple reaction monitoring–mass spectrometry

QC: quality control

LOD: Limit of detection

LOQ: Limit of quantification

LLOQ: lower limit of quantification

ULOQ: upper limit of quantification

AUROC: area under the receiver operating characteristic

ACN: Acetonitrile

FA: Formic acid

ABC: Ammonium bicarbonate

DTT: Dithiothreitol

IAA: Iodoacetamid

QqQ: Triple quadrupole

LC: Liquid chromatography

Glu: Glutamic acid

Gla: Gamma-carboxy glutamic acid

SIS: Stable isotope-labeled internal standard

CV: Coefficient of variation

I. INTRODUCTION

Liver cancer is the fifth most common cancer and the second cause of cancer-related death. The incidence rate and mortality rate were 10.1 and 9.5 per 100,000 persons in 2012, respectively (1). Hepatocellular carcinoma (HCC) is the major histologic type among primary liver cancers occurring worldwide, accounting for 70% to 85% of the total burden. Early detection of HCC increases the potential for curative treatment and improves prognosis. Several methods developed for the diagnosis of HCC including evaluation of blood markers have been clinically tested.

Alpha fetoprotein (AFP) is a glycoprotein with a complex, single asparagine-linked sugar structure that has been widely used as a blood biomarker for HCC (1). Approximately 25 sugar structures have been identified at a single site in AFP, primarily by mass spectrometry (3). Among them, the lens culinaris agglutinin-reactive fraction of alpha-fetoprotein (AFP-L3), a glycoform with core fucosylation, has been known as a HCC-specific biomarker. In clinical practice, AFP-L3% is defined as the ratio of the AFP-L3 fraction to the total AFP concentration (4). Also, protein induced by vitamin K absence or antagonist-II (PIVKA-II), an abnormal prothrombin induced by vitamin K absence or antagonist which lacks the capability of reciprocal action with other coagulation factors, is another serum marker used for both surveillance of at-risk patients and HCC diagnosis. (1). In addition to AFP, AFP-L3% and PIVKA-II, a number of other biomarkers have been evaluated for early diagnosis in research phase. These include glypican-3, squamous cell carcinoma antigen, cytokeratin 19 and golgi protein-73, but there have not been strong

consensus for performance of these biomarkers (1). Although more and more research is under development of novel biomarkers, the study on whether certain biomarkers can be utilized in clinical are still in real worldwide demand (2).

In this study, I developed the method for quantifying AFP, AFP-L3 and PIVKA-II in serum by capillary-flow liquid chromatography, interfaced with an MRM-MS-based assay that can overcome the low analytical sensitivity of the existing method. Currently, in clinical practice, liquid-phase binding assay (LiBA) is the clinical standard for measuring AFP, AFP-L3%, and PIVKA-II. In particular, the micro-total assay system (μTASWako™i-30), which obtained clearance from the US Food and Drug Administration (FDA) for *in vitro* diagnostic use in February 2011 (7, 8), is the preferred system of most major reference laboratories in the US. The existing method for quantifying AFP-L3 in serum lack the analytical sensitivity to make accurate diagnoses (5, 9). However, compared with existing method, the MRM-MS assay was able to quantify three markers with wider dynamic ranges. The quantifiable concentrations for MRM-MSassay range from 0.051 to 4000 ng/mL for AFP; from 0.132 to 100% for AFP-L3%; from 1.28 to 100000 ng/mL for PIVKA-II. Using 400 serum samples (100 chronic hepatitis, 100 liver cirrhosis, 200 HCC), the performance of the MRM-MS assay was compared with that of LiBA, conventional method, demonstrating the MRM-MS is superior to the conventional method (LiBA) of distinguishing HCC from noncancer patients for early diagnosis.

Subsequently, I validated analytical performance of the developed MRM-MS assay for AFP, AFP-L3 and PIVKA-II. With regard to quantification of analyte, the selective, sensitive and reproducible method is critical for the assay development.

Meanwhile, laboratory-to-laboratory variability for the same analyte measured by LC-MS, has been observed as evidenced by results from external quality assessment programs and published reports (3). To date, few organized efforts have been made to harmonize LC-MS methods across different clinical laboratories. Until recently, there existed minimal guidance on the use of LC-MS for clinical diagnostics. Given the growing use of LC-MS in clinical laboratories and the wider range of people who use the technology, there is an increased need for robustness and harmonization of LC-MS methods. Thus, MRM-MS assay was fully validated per integrated guidelines by the US FDA, European Medicines Agency (EMA), Korea FDA (KFDA), and Clinical and Laboratory Standards Institute (CLSI). Through the method validation, I confirmed that the MS-based assay for AFP, AFP-L3, and PIVKA-II met the global consensus standards for health care testing.

In conclusion, throughout this study, to establish its methodological traceability, laboratory-developed test (LDT) was characterized extensively with respect to the analytical quality requirements for quantitative protein biomarker assays.

II. MATERIALS AND METHODS

2.1. Clinical specimen

All blood samples were incubated in BD Vacutainer® blood collection tubes for 30 min (clotting time, at room temperature) and centrifuged at 1200 g for 20 min at room temperature. Supernatant aliquots (300 μL) were stored in plain tubes at -80 °C until analysis. All serum samples were collected at a single institution from 2008 to 2014 per standard operating procedures (16). The use of human serum samples was approved by the institutional review board of Seoul National Hospital (IRB No. 0506-150-005). A total of 100 samples from patients with chronic hepatitis (71 men, 29 women; median age 56 y, range 29–69 y), 100 samples from patients with liver cirrhosis (38 men, 62 women; median age 58 y, range 34–78 y), and 200 samples from patients with hepatocellular carcinoma (HCC) (166 men, 34 women; median age 59 y, range 38–86 y) were obtained (Table 1).

Table 1. Demographics and clinical characteristics of the study population (*N* = 400)

		Hepatitis (*N* = 100)	Liver cirrhosis (*N* = 100)	HCC (*N* = 200)	P-value
Age (years)					
	mean ± SD	54.2 ± 11.1	57.5 ± 9.8	59.4 ± 10.6	0.109
Gender					
	M/F	71/29	38/62	166/34	<0.0001
Etiology					
	HBV/HCV/Alcohol/Unknown	100/0/0/0	80/8/12/0	158/16/10/16	<0.0001
ECOG					
	0/1	100/0	100/0	184/16	<0.0001
Tumor size (cm)					
	< 5			169 (84.5%)	
	≥ 5			31 (15.5%)	
Tumor number					
	≤ 3			182 (91.0%)	
	4-9			8 (4.0%)	
	≥ 10			10 (5.0%)	
Vascular invasion					
	No			181 (90.5%)	
	Yes			19 (9.5%)	
Lymph node					
	No			195 (97.5%)	
	Yes			5 (2.5%)	
Metastasis					
	No			198 (99.0%)	
	Yes			2 (1.0%)	

HCC, hepatocellular carcinoma; HBsAg, hepatitis B surface antigen; anti-HCV, antibody against hepatitis C virus; ECOG, eastern cooperative oncology group.

2.2. Chemical and Reagents

Protein analogs that contained unlabeled and labeled (13C15N Arg/Lys-labeled) recombinant AFP protein were purchased from Sino Biological Inc. (Beijing, China) and Origene Technologies Inc. (Rockvile, MD, USA), respectively. The absence of sugar residues that were attached to the protein analogs was confirmed by mass spectrometry. Amino acid analysis was performed by AAA Service Laboratory to confirm the quality of the content before use (purity > 95%). Mouse monoclonal anti-AFP antibody was purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA) and raised against amino acids 171–310 of AFP, which contains the central region of the glycosylation site. For immunoaffinity enrichment, anti-AFP antibody was covalently coupled to Dynabeads Protein G (immunoprecipitation kit, Invitrogen, Carlsbad, CA, USA). For the AFP-L3 fractionation, lens culinaris agglutinin (LCA) lectin-MagneZoom was purchased from bioWORLD (Dublin, OH, USA). Formic acid, dithiothreitol, iodoacetamide, ammonium bicarbonate, phosphate-buffered saline, methyl α-D-mannopyranoside, sodium chloride, manganase chloride, and calcium chloride were purchased from Sigma-Aldrich (St. Louis, MO, USA). RapiGest was purchased from Waters (Manchester, UK), and sequencing-grade modified trypsin was obtained from Promega (Madison, WI, USA). Peptide-N-glycosidase-F (PNGase-F) was purchased from New England BioLabs Inc. (Beverly, MA, USA). Protein LoBind tubes were purchased from Eppendorf (Westbury, NY, USA).

2.3. AFP enrichment and AFP-L3 fractionation

I bound 10 μL of clear Dynabeads (30 μg/μL) to 2 μL of monoclonal anti-AFP antibody (0.5 μg/μL), diluted it with 20 μL antibody binding and washing buffer (manufacturer's buffer, composition confidential), and incubated it at room temperature for 10 min. The supernatant was discarded twice. The Dynabeads-Ab complex was added into a serum sample (200 μL) that contained 10 μL of stable isotope-labeled internal standard protein analog (38.4 ng/μL) and incubated with rotation at room temperature for 30 min. The supernatant was discarded twice and subsequently washed with 50 μL washing buffer (manufacturer's buffer, composition confidential). Ten microliters of 5% acetic acid was used as the elution buffer after incubation with rotation at room temperature for 2 min.

The supernatant (containing enriched AFP protein) was transferred to a new tube and diluted with 10 μL of binding buffer (20 mM Tris-HCl, 0.5 M NaCl, 1 mM MnCl2, and 1 mM CaCl2). Then, 20 μL of magnetic-LCA lectin (25 μg/μL) was added and incubated at room temperature with rotation for 30 min. After the incubation, the magnetic beads were washed twice with 10 μL of binding buffer. Bound analyte (the AFP-L3 fraction) was eluted with 20 μL of elution buffer (100 mM α-methyl mannopyranoside, 1X phosphate-buffered saline) after 15 min of mixing with rotation at room temperature. The eluate (bound analyte: AFP-L3 fraction) and supernatant (unbound analytes: AFP-L1, L2 fraction) were transferred to a new tube. All sample preparation steps were performed using a magnetic particle concentrator (Invitrogen, Carlsbad, CA, USA).

2.4. Deglycosylation and digestion

Prior to digestion, the AFP-L3 fraction was deglycosylated with 2 μL of peptide-N-glycosidase-F (PNGase-F, 500000 units/mL) at 37 ℃ for 4 h. The AFP-L1, L2 fraction was reacted with 50 mM ammonium bicarbonate, pH 8.0, instead of PNGase-F. A stock solution was prepared by adding 0.2% RapiGest surfactant and 20 mM dithiothreitol in 100 mM ammonium bicarbonate, pH 8.0. Twenty microliters of stock solution was added to the sample and vortexed before being incubated at 60 °C for 60 min. After the incubation, the sample was alkylated with 10 μL of 100 mM iodoacetamide, vortexed, and incubated in the dark at room temperature for 30 min. Then, 40 μL of trypsin solution (0.1 μg/μL) was added to the sample. The sample was vortexed and incubated at 37 °C for 4 h.

The digestion was completed by adding 10 μL of 10% formic acid to the sample. The sample was centrifuged at 15000 rpm at 4 °C for 60 min to remove insoluble chemicals, such as byproducts of RapiGest surfactant. The supernatant from the AFP-L3 and AFP-L1, L2 fractions was transferred to a new tube. The overall time that was required for the sample preparation was approximately 12 h. All sample preparation steps were performed on a Thermomixer C (Eppendorf, Westbury, NY, USA). The digested samples were moved to an autosampler for MRM-MS analysis.

2.5. MRM-MS Analysis

Chromatographic separation of GYQELLEK (representing AFP) and VDFTEIQK (representing AFP-L3) ERECVEETCSY (representing PIVKA-II) from matrix

components before MRM-MS analysis was performed on a fully automated online 1260 Capillary-flow liquid chromatography (LC) system (Agilent Technologies, Santa Clara, CA, USA). The sample compartment of the autosampler was set to 4 °C, and the LC separation was conducted at 40 °C. Sample cleanup was performed using a guard column (2.1 × 15.0 mm, 1.8 μm, 80 Å), and peptides were separated on an analytical column (0.5 × 35.0 mm, 3.5 μm, 80 Å) (both columns from Agilent Technologies, Santa Clara, CA, USA). The 2 columns in the LC system were operated in parallel, facilitating high throughput and using the same solvents: A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile).

Ten microliters of the digested sample were injected onto a guard column with the effluent directed to waste, at 40 μL/min for 1 min in 10% solvent B. After the valves were switched, sample flow was directed from the guard to the analytical column. After switching the valve, 10% solvent B was run at a flow rate of 40 μL/min for 1 min. Bound peptides were then eluted from the analytical column on a linear gradient of 10% to 60% solvent B over 5 min at 40 μL/min. The column wash was held at 90% solvent B for 1 min at 40 μL/min and then equilibrated at 10% solvent B for 4 min. After 1 min of equilibration, the switching valve was returned to its original position, and reconditioning occurred simultaneously in the guard and analytical columns. The total injection-to-injection run time, including sample cleanup and re-equilibration, was 12 min. The autosampler injector needle and tubing were washed after each sample injection with aqueous 50% acetonitrile solution.

Quantitative analysis was performed on an Agilent 6490 triple quadrupole

(QqQ) mass spectrometer with a Jetstream electrospray source (Agilent Technologies, Santa Clara, CA, USA), operated in positive ion, multiple reaction monitoring (MRM) mode. The source parameters were as follows: gas temperature, 250 °C; gas flow, 15 L/min; nebulizer, 30 psi; sheath gas temperature, 350 °C; sheath gas flow, 12 L/min. The delta electron multiplier voltage (EMV) was set to 200 V, and the cell accelerator voltage and fragment voltage were 5 V and 380 V, respectively. The dwell time and cycle time of the mass transitions were 66 millisecond (ms) and 2502 ms, respectively. The Q1 quadrupole and Q3 quadrupole were set to unit resolution: 0.7 Da at half height.

2.6. Liquid-phase binding assay (LiBA)

The μTAS autoanalyzer (Wako Pure Chemical Industries) is a FDA-clearance device for diagnosing HCC by measuring AFP and AFP-L3 concentrations. A sample load of 100 μL was analyzed for 9 min with a 2 min interval between each sample. The AFP-L3 concentration was calculated automatically as a percentage of total AFP and printed out. The quantifiable ranges of AFP and AFP-L3 were 0.3–4000.0 ng/mL and 0.5% to 99.5%, respectively, using a 2-point calibrator. All serum samples were measured following the manufacturer's instructions.

2.7. Data Analysis

Skyline, ver. 3.7 (MacCoss Lab, University of Washington, USA) was used to import and align all MRM-MS raw data files and quantitative features. The absolute abundance of transitions was determined, based on the normalized peak areas of the

purified internal standard. The data points on the peak chromatogram were smoothed by Savitzky-Golay method (1). All data were analyzed repeatedly using the mean (or median) value. Statistical analysis (Mann-Whitney test, area under the receiver operating curve, correlation analysis) and visualizations were performed using MedCalc, ver. 12.7 (MedCalc, Mariakerke, Belgium), IBM SPSS, ver. 19.0 (SPSS Inc., Chicago, IL, USA), and Graph Pad Prism, ver. 6.0 (Graph PAD, San Diego, CA, USA).

2.8. Calculation of AFP-L3 concentration by MRM-MS assay

AFP-L3% concentration was calculated as follows: After MRM-MS analysis, the peak area ratios (unlabeled to labeled) of the glycopeptide (VDFTEIQK), which served as a surrogate peptide for AFP-L3, were obtained. The log2-transformed peak area ratio was used for back-calculation, based on the forward response curve. The back-calculated value represented the log2-transformed ratio of unlabeled to labeled peptide in concentration units (ng/mL), which was then converted to the original value. The original value was multiplied by the concentration of spiked stable isotope-labeled internal standard protein analog (19.2 ng). Then, the concentration (ng/mL) of endogenous AFP-L3 was obtained. Similarly, the concentration (ng/mL) of total AFP was obtained, based on the peak area ratio of nonglycopeptide (GYQELLEK). Finally, AFP-L3% was calculated by dividing the AFP-L3 concentration that was obtained with the VDFFTEIQK peptide by the total AFP concentration that was obtained with the GYQELLEK peptide and multiplying the result by 100.

2.9. Analytical method validation

The analytical method validation experiments were designed to meet the requirements (validation practices) of 4 sets of guidelines [US Food and Drug Administration (FDA) (2), European Medicines Agency (EMA) (3), Korea FDA (KFDA) (4), and Clinical and Laboratory Standards Institute (CLSI)] (5-19). The acceptance criteria (performance specification) for all method validation items were based on the guidance documents and other published whitepapers (20-30). These items covered such aspects as calibration curve, analytical specificity (selectivity or interference), analytical sensitivity, carryover, precision, recovery of assay, matrix effect, recovery of immunoprecipitation, dilution integrity, stability, reproducibility, and quality control (QC) of samples and frequency. A schematic diagram of the validation of the analytical method and the optimized analytical sequence and schedule are provided in Figure 1and Table 2, respectively.

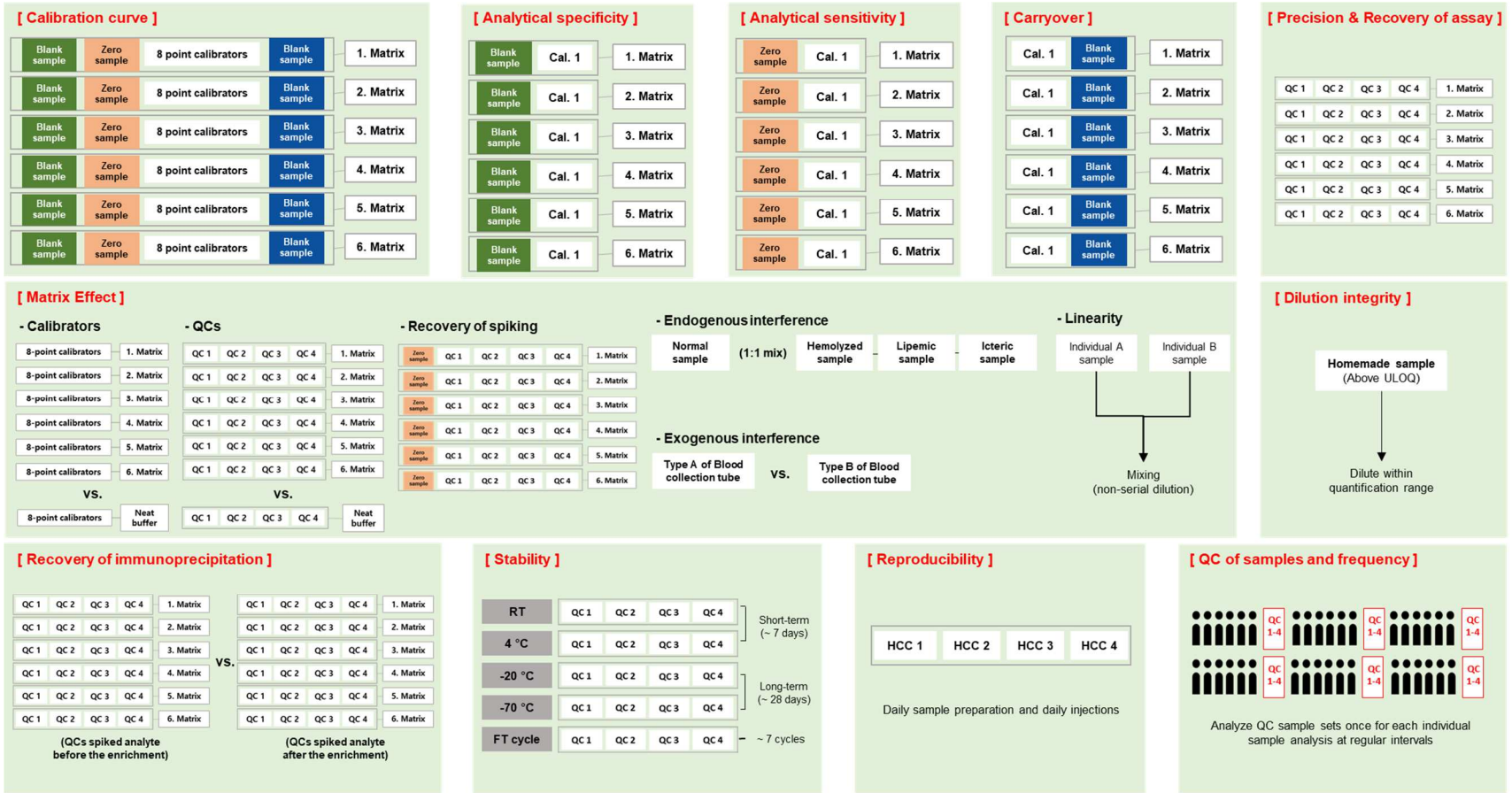
[Calibration curve]
Blank sample
Zero sample
8 point calibrators
Blank sample
1. Matrix
2. Matrix
3. Matrix
4. Matrix
5. Matrix
6. Matrix
[Analytical specificity]
Blank sample
Cal. 1
[Analytical sensitivity]
Zero sample
Cal. 1
[Carryover]
Cal. 1
Blank sample
[Precision & Recovery of assay]
QC 1
QC 2
QC 3
QC 4
[Matrix Effect]
- Calibrators
8-point calibrators
vs.
Neat buffer
- QCs
- Recovery of spiking
- Endogenous interference
Normal sample
(1:1 mix)
Hemolyzed sample
Lipemic sample
Icteric sample
- Exogenous interference
Type A of Blood collection tube
vs.
Type B of Blood collection tube
- Linearity
Individual A sample
Individual B sample
Mixing
(non-serial dilution)
[Dilution integrity]
Homemade sample
(Above ULOQ)
Dilute within
quantification range
[Recovery of immunoprecipitation]
vs.
(QCs spiked analyte before the enrichment)
(QCs spiked analyte after the enrichment)
[Stability]
RT
4 °C
-20 °C
-70 °C
FT cycle
Short-term
(~ 7 days)
Long-term
(~ 28 days)
~ 7 cycles
[Reproducibility]
HCC 1
HCC 2
HCC 3
HCC 4
Daily sample preparation and daily injections
[QC of samples and frequency]
QC 1-4
Analyze QC sample sets once for each individual sample analysis at regular intervals

Figure 1. Schematic diagram of analytical method validation by MRM-MS assay.

Sample configuration of each evaluation item in the analytical method validation study. The method validation study consisted of 12 items: calibration curve, analytical specificity (selectivity or interference), analytical sensitivity, carryover, precision, recovery of assay, matrix effect, dilution integrity, recovery of immunoprecipitation, stability, reproducibility, and quality control (QC) of samples and frequency.

Table 2. Optimized analytical sequence and schedule for analytical method validation by MRM-MS assay.

Analysis order	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	Evaluation or Analysis
Calibration curve	■																												**Calibration curves, Analytical specificity, Analytical sensitivity, and Carryover evaluation**
Individual sample		■																											**Individual sample analysis**
QC of samples and frequency		■																											**Reliability individual sample analysis**
Stability																													
0 d			■																										**Stability evaluation**
1 d (short & long-term)				■																									
2 d (short & long-term)					■																								
3 d (short & long-term)						■																							
4 d (short-term)							■																						
5 d (short-term)								■																					
6 d (short-term)									■																				
7 d (short & long-term)										■																			
14 d (long-term)																	■												
21 d (long-term)																								■					
28 d (long-term)																												■	

Precision & Recovery of assay																													
1 d											■																		**Evaluation for Precision and Recovery of assay**
2 d												■																	
3 d													■																
4 d														■															
5 d															■														
6 d																■													
Reproducibility																													
1 batch																	■												**Reproducibility evaluation**
2 batch																		■											
3 batch																			■										
4 batch																				■									
5 batch																					■								
6 batch																						■							
Matrix effect																													
Calibrators, QCs, recovery of spiking, matrix mixing, endogenous, exogenous																								■					**Matrix effect evaluation**
Recovery of immunoprecipitation																									■				**Evaluation for Recovery of immunoprecipitation**
Dilution integrity																										■			**Dilution integrity evaluation**

2.9.1. Calibration curve

Due to the presence of analytes in the serum matrix, it was not possible to calculate the lower limit of quantification (LLOQ) accurately through the forward calibration curve. Therefore, I generated a reverse calibration curve to accurately and reproducibly calculate the LLOQ by circumventing the endogenous analytes. The reverse calibration curve was constructed by spiking equal amounts of the unlabeled protein analog with varying amounts of the labeled protein analog at 8 concentrations. Conversely, the forward calibration curve was generated by spiking equal amounts of the labeled protein analog and varying amounts of the unlabeled protein analog at the same 8 concentrations as with the reverse calibration curve. Blank samples (matrix only, no internal standard) and zero samples (matrix spiked with internal standard) of each serum matrix were analyzed.

Reverse and forward calibration curves were generated using each of the 6 chronic hepatitis serum samples as the matrix. In both of the calibration curves, the internal standard protein analog was spiked constantly at the middle point (GYQELLEK, VDFTEIQK; 19.2 ng/mL, and ERECVEETCSY;480 ng/mL) over the linear range (GYQELLEK and VDFTEIQK; 0.051–4,000 ng/mL, ERECVEETCSY; 1.28–100000 ng/mL) to reduce fluctuations in peak area ratio. All calibrators of the 6 matrices were prepared daily and analyzed in 3 replicates per day over 6 d. The equations that were used to quantify the analytes were generated by averaging the parameters (slope, intercept) of the 6 equations that were obtained from the 6 matrices. The reverse calibration curve was used to calculate the concentrations of the samples, which contained stable isotope-unlabeled internal

standard protein analog. Conversely, the forward calibration curve was used to calculate the concentrations of patient samples that contained stable isotope-labeled internal standard protein analog. The upper limit of quantification (ULOQ) was defined as the highest concentration in the reverse calibration curve that met the following criteria: precision (CV < 20%), and recovery of assay (within ± 15% of target). In this experiment, the ULOQs of AFP, AFP-L3 were 4000 ng/mL and ULOQ of PIVKA-II was 100000 ng/mL. Due to financial constraints and the high cost of protein analogs, I were unable to evaluate concentrations above 4000 ng/mL or 100000 ng/mL.

2.9.2. Analytical specificity (selectivity or interference)

Analytical specificity was calculated by comparing the concentrations of endogenous analyte and internal standard in the blank samples, with those of the LLOQ (calibrator 1). To elaborate, interference of the analyte was calculated as the percentage of the peak area of labeled peptide in the blank samples with respect to the peak area of labeled peptide in the samples at the LLOQ. For the internal standard, interference was calculated as the percentage of the peak area of unlabeled peptide in the blank samples with respect to the peak area of the unlabeled peptide in the samples at the LLOQ. Pairs of interference were calculated across all 6 matrices in triplicate per day over 6 d.

2.9.3. Analytical sensitivity

Individual chronic hepatitis serum from 6 patients were used as the matrix for determining the LLOQ. The 6 matrices were prepared daily and analyzed in triplicate each day over 6 d. The LLOQ was defined as the lowest concentration on the reverse calibration curve (calibrator 1) that could be measured with a range of precision of < 20% and a recovery of assay of 80% to 120% of target concentrations with a S/N of at least 5. Notably, the CLSI guidelines are stricter (recovery of assay < ±15%, S/N > 20) than other criteria.

2.9.4. Carryover

Carryover was evaluated similarly as analytical specificity, except that the blank samples were analyzed sequentially after the sample at the ULOQ (calibrator 8). Six blank samples were evaluated from several chronic hepatitis serum samples. The 6 matrices were prepared daily and analyzed in triplicate each day for 6 d. Carryover was deemed to be satisfactory if it met the following requirements: the peak area of labeled peptide (analyte) in the blank samples were less than 20% of that in the samples at the LLOQ (calibrator 1) and the unlabeled peptide (internal standard) peak area in the blank samples were less than 5% of that in the samples at the LLOQ. Therefore, the carryover limit, defined as the highest concentration that carries over to subsequent samples, should be determined. EMA and CLSI guidelines recommend that a blank sample be analyzed immediately following a ULOQ sample to ensure that carryover will not affect subsequent specimens.

2.9.5. Precision and Recovery of assay

Precision and recovery of assay were assessed by analyzing 4 quality control (QC) samples. Four QC samples were prepared by spiking stable isotope-labeled protein analog into the chronic hepatitis serum as the matrix at the following concentrations: QC 1 (LLOQ, 0.051 ng/mL), QC 2 (3 × LLOQ, 0.154 ng/mL), QC 3 [(LLOQ + ULOQ) / 2000 ng/mL], QC 4 (0.9 × ULOQ, 3600 ng/mL) for AFP and AFP-L3, and QC 1 (LLOQ, 1.28 ng/mL), QC 2 (3 × LLOQ, 3.84 ng/mL), QC 3 [(LLOQ + ULOQ)/2, 50,001 ng/mL], QC 4 (0.9 × ULOQ, 90,000 ng/mL) for PIVKA-II,. Each QC sample was analyzed in 6 replicates per day for 6 d. The CV was used to report the intra-assay and inter-assay precision. The CV values were expected to be less than 15% for all concentrations except for QC 1, which had a concentration that was equivalent to the LLOQ (< 20%), to meet the requirements of the guidelines. The recovery of intra-assay and inter-assay relative to target should also be within ± 15% for all concentrations except for QC 1 (± 20%).

The precision and recovery of intra-assay were calculated as described in the following segment. First, 6 replicates of the QC 1 sample were analyzed each day for 6 d. The CV values and concentrations of each of the 6 replicates on a given day were calculated and averaged. Subsequently, these mean CV values and concentrations for each day were averaged to calculate the overall precision and recovery of intra-assay. The first replicate of each day was averaged to calculate the precision and recovery of inter-assay. This process was repeated for all remaining QC samples.

2.9.6. Matrix effect

The matrix effect was evaluated by preparing 8 calibrators and 4 QC samples from each of the 6 chronic hepatitis serum as matrices. The calibrators were analyzed in triplicate. In contrast, the QC samples were analyzed in 6 replicates for a given QC sample in a particular matrix. This step was repeated for all samples each day over 6 d. The matrix effects were calculated by comparing the peak area ratio (labeled to unlabeled) of the analyte in the serum matrix with that in neat solution (antibody binding and washing buffer) at the same concentrations. The resulting ratio was expressed as %matrix effect.

The recovery of spiking was calculated for the 4 QC samples, based on the concentrations in the zero samples. Recovery of spiking is used to determine whether a systemic shift occurs in the analytical signal of an analyte due to a matrix effect. The concentration of QC samples was determined for the spiked (a) and unspiked zero samples (b). Based on the expected (theoretical) concentration of the spiked zero samples in QC samples (c), the percent recovery of spiking is calculated as (a-b)/c ×100.

The matrix effect was evaluated in a matrix mixing experiment. If significant sample-specific matrix effects were present, the measured concentrations will not agree with the expected concentrations. To confirm this, 2 HCC serum samples were mixed at varying proportions (100:0, 80:20, 60:40, 40:60, 20:80, and 0:100) to evaluate the linearity of the patient samples. HCC serum samples were selected, based on the concentrations as measured by LiBA. Serum A (low AFP) and

B (high AFP) and C (low AFP-L3) and D (high AFP-L3) were used to analyze the linearity of GYQELLEK and VDFTEIQK peptides, respectively. All samples were analyzed in 6 replicates in 1 d.

Endogenous compounds like hemolysis, lipemia, and icterus (problem samples) in blood can produce a matrix effect that suppresses or enhances ionization. To assess the extent to which such abnormalities affect the quantitative results of patient samples, concentrations of AFP, AFP-L3 and PIVKA-II were measured in lipemic, hemolyzed, and icteric samples, mixed at a 1:1 ratio with “normal” serum sample. All samples were analyzed in 6 replicates per day in 1 d.

I evaluated exogenous compounds using 2 types of tubes [serum separation tubes (SSTs, serum) and tri-potassium-ethylenediaminetetraacetic acid (K3-EDTA, plasma] to store blood from 6 patients with HCC. The 6 samples were selected, based on an even distribution of concentrations in the quantification range. All samples were analyzed in 6 replicates in 1 d. The concentrations of AFP, AFP-L3, PIVKA-II in serum and plasma were determined, respectively.

2.9.7. Recovery of immunoprecipiation

The recovery after enrichment by immunoprecipitation was determined in 4 QC samples of 6 chronic hepatitis serum matrices. Stable isotope-unlabeled and labeled protein analogs were added before the AFP or PIVKA-II enrichment step in 1 set of samples, whereas it was added to another set after enrichment step. Each sample was

analyzed in 6 replicates per day for 6 d. Recovery of immunoprecipitation was expressed as the relative recovery in peak area ratio of the enriched versus unenriched samples. Recovery measurements for AFp-L3 fractionation were unavailable due to a lack of standard material for the AFP-L3 glycoform.

2.9.8. Dilution integrity

The effect of sample dilution on precision and recovery of assay was assessed by spiking the chronic hepatitis sample (blank sample) with stable isotope-labeled protein analog at 8000 ng/mL (AFP), 300000 ng/mL (PIVKA-II) and measuring the resulting concentrations at various dilutions, established by diluting a blank sample by 5 factors (AFP; 5, 150, 500, 2500, 20000, PIVKA-II; 200, 400, 2000, 10000) that were within the dynamic range of the assay. Each sample was diluted separately instead of serially. Each sample was analyzed in 6 replicates in 1 d. The precision and recovery of assay needed to be within 15% to meet the guideline standards.

2.9.9. Stability

The stability in storage was evaluated to determine acceptable storage conditions. Stability studies were conducted with 4 QC samples under various storage conditions. To assess short-term storage stability, the samples were stored on the benchtop at room temperature and in an autosampler at 4 °C for up to 7 d. Long-term stability was assessed by storing the samples in a freezer at -20 °C and -70 °C for up to 28 d.

Freeze-thaw stability was also assessed in the 4 QC samples. Aliquots were subjected to 0 to 7 freeze-thaw cycles and then analyzed in triplicate. Sample stability was calculated as the mean difference from the baseline value (0 d or 0 cycle), with 85% to 115% as the acceptable range.

2.9.10. Reproducibility

The reproducibility of the entire MRM-MS assay workflow was evaluated with 4 HCC serum samples, which were selected based on an even distribution of concentrations of AFP, AFP-L3 and PIVKA-II within the quantifiable range. The serum samples were prepared fresh each day (16 h apart), and the 4 samples were analyzed daily in 6 replicates over 6 d.

2.9.11. Quality control (QC) of samples and frequency

To ensure that the patient samples were analyzed under the proper conditions, 4 QC samples were analyzed after every 20 individual samples. A total of 400 individual samples were analyzed, with 20 sets of QC runs (5% of total patient size for each QC concentration), meeting the guideline requirement that QC runs constitute at least 5% of the total patient size.

Each individual sample was analyzed once over 20 d. For the 4 QC measurements to be valid, over two-thirds of all QC samples needed to have concentrations less than 15% of the nominal concentration. In addition, at least 50%

of each QC measurements must be within ±15% of the nominal concentration to be valid.

III. RESULTS

3.1. Development of the assay based on the mass spectrometry

3.1.1. Confirmation of the surrogate peptide of AFP and AFP-L3

For the selection of the surrogate nonglycopeptide, Skyline software was used to generate a list of all possible b- and y-series fragment ions (1+ or 2+ ion charge state) for the 2+ or 3+ precursor ion charge state, spanning the m/z range from 300 to 1400. As shown in Figure 2A, in 37 tryptic peptides of AFP, there were 26 predicted unique tryptic peptides with a useful length of 6–30 amino acids. Among these candidates, those with potential ragged ends (-KK, -RR, -KR, and -RK) were removed. The first 25 amino acids counting from the N-terminus of the protein were not considered when selecting candidates due to the possibility of containing a signal peptide sequence. Several peptides with (n-term) cysteine (Cys), methionine (Met), histidine (His), an NxS/T motif, and pyroglutamine (pyroGln), which are susceptible to in vivo and in vitro modifications that affect their mass-to-charge (m/z) ratios, were also eliminated. In addition, peptides with RP/KP, KxK/R, or RxK/R were eliminated due to low digestion efficiency. Ultimately, 5 nonglycopeptides remained after these criteria were applied. The stable isotope-unlabeled protein analog (1000 ng) was digested and analyzed in triplicate. Among 5 peptides, the GYQELLEK peptide had the highest intensity when measured by MRM-MS assay and was selected as a surrogate peptide of AFP (Figure 2B).

Based on the sequence in the Uniprot database (http://www.uniprot.org),

the AFP glycopeptide has a single N-glycosylation site. According to a public glycan database (UniCarbKB, http://unicarbkb.org), AFP protein has 25 glycan structures (22 biantennery and 3 triantennary N-linked oligosaccharides) through a combination of 5 sugar residues (N-acetylglucosamine, Mannose, Galactose, N-acetylneuramic acid, and Fucose) at amino acid Asn-251. To quantify the heterogeneous fucose-attached AFP-L3 moieties simultaneously, we analyzed the PNGase-F-treated deglycopeptide, in which the sugar moiety is cleaved and an asparagine (Asn, N) residue in the NxS/T motif is changed into aspartic acid (Asp, D).

The unit resolution was set to 0.7 Da for both the Q1 quadrupole and Q3 quadruple, and a high signal-to-noise ratio (S/N) was obtained with sufficient selectivity for typical samples. For the AFP-L3 measurements in the patient samples by the MRM-MS assay, 2 similar types of peptide sequences (deglycosylated form, VDFTEIQK sequence as endogenous AFP-L3; original form, VNFTEIQK sequence, as internal standard) should be quantified. The 2 similar sequences were not distinguished by the unit-unit resolution, because the difference in precursor ions at the double-charge was 0.4 Da, and the selected product ions (y6) for quantification were also identical.

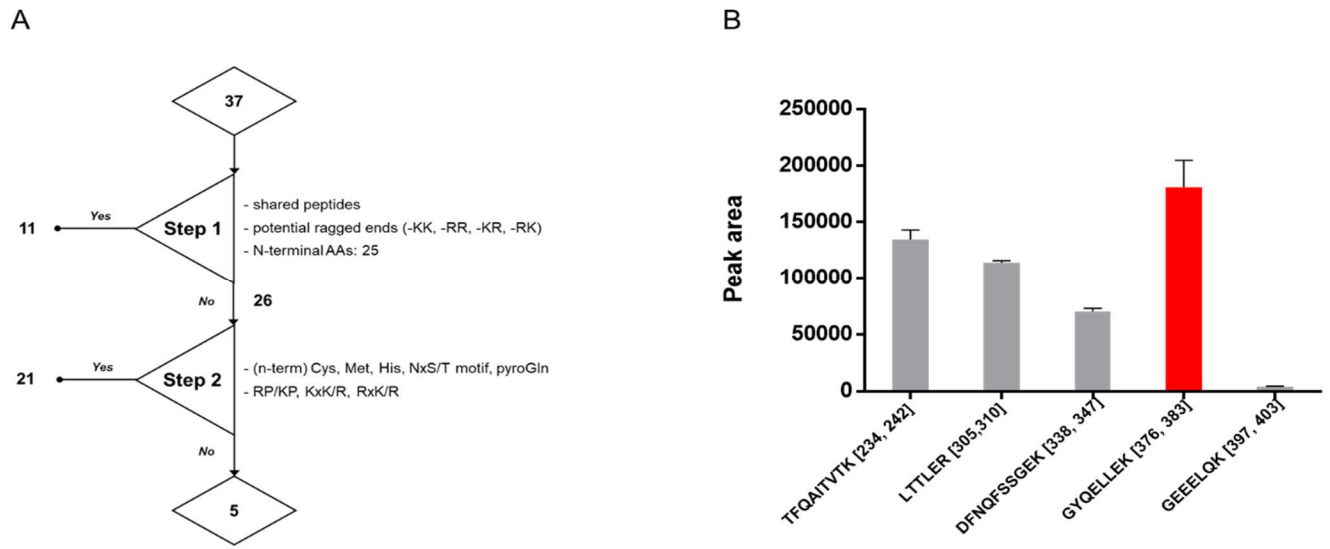

Figure 2. Selection of the surrogate nonglycopeptide of AFP.

In step 1, 26 unique peptides were selected. These peptides did not contain potential ragged ends or a signal peptide sequences. Of these peptides, 21 were excluded, because they harbored amino acids that hindered the detection with high intensity by mass spectrometer (step 2). Ultimately, 5 peptides were selected as surrogate peptide candidates (A). The stable isotope-unlabeled protein analog (1000 ng) was digested and analyzed in triplicate. Finally, GYQELLEK peptide (red) was detected to have the highest intensity by MRM-MS assay (B).

3.1.2. Modeling the surrogate peptide for representing PIVKA-II

The first step in developing an MRM-MS assay is to determine the surrogate peptide for quantitative analysis. Prothrombin is converted to its active form after the 10 Glu residues at positions 6, 7, 14, 16, 19, 20, 25, 26, 29, and 32 in the N-terminal Gla domain are γ-carboxylated to Gla by γ-glutamylcarboxylase (Figure 3A) (4). In patients with HCC, γ-glutamylcarboxylase is impaired; thus, the 10 Glu residues can not be γ-carboxylated completely to Gla. The PIVKAII variants that are preferentially synthesized in HCC patients bear less than 4 Gla residues, whereas those in benign liver diseases have more than 5 (5). To diagnose HCC patients, I wanted to select the most representative surrogate peptide that contained Glu residues in the Gla domain. Initially, the full sequence of prothrombin was imported into Skyline, in which proteolytic digestion was performed *in silico*. Skyline supports various enzymes, and I began to develop the peptide with trypsin and chymotrypsin. As a result, 2 tryptic peptides (ANTFLEEVR and ECVEETCSYEEAFEALESSTATDVFWAK) and 3 chymotryptic peptides (EEVRKGNL, ERECVEETCSY, and ESSTATDVF).

Gamma-carboxylated forms of the selected 2 tryptic peptides and 3 chymotryptic peptides were modeled *in silico* using Skyline. As a result, 68 and 21 of all possible combinations of gamma-carboxylated forms were modeled for the tryptic and chymotryptic peptides, respectively, by changing each Gla to Glu (Table 3). MRM-MS analysis was then performed, targeting all 89 peptide combinations, using serum from 3 patients, containing high, medium, and low PIVKAII levels by

immunoassay. Consequently, none of the 68 modified tryptic peptides produced a notable signal at the predicted mass values, perhaps because 4 modeled peptides from the ANTFLEEVR sequence had ragged ends (-RR and -RK) that lowered the efficiency of trypsin, whereas 64 modeled peptides from the ECVEETCSYEEAFEALESSTATDVFWAK sequence were long and hydrophobic (Figure 1-2 A).

Ultimately, I detected only 1 chymotryptic peptide by MRM-MS, ERECVEETCSY, in which none of the Glu residues was converted to Gla (Figure 3B). The intensity of this peptide was consistent between the 3 PIVKA-II levels in the patient serum samples (Figure 4). Based on these findings, this chymotryptic peptide (ERECVEETCSY) satisfied the following criteria: (1) length between 6 and 30 amino acids; (2) mass greater than 800 m/z (to facilitate detection by electrospray ionization-MS); and (3) no Met residues (to avoid chemical modifications). Also, the chymotryptic peptide, the sequence of which was related to only 1 protein in the Uniprot/Swiss-Prot database, was considered a uniquely representative proteolytic sequence of PIVKA-II. Consequently, I determined the proteolytic peptide ERECVEETCSY to be a surrogate peptide that could be used to quantify PIVKA-II levels in human blood.

The parameters of the 3 peptides (GYQELLEK, VDFTEIQK, ERECVEETCSY) and their transitions are presented in Table 4.

Table 3. List of candidate peptides for selecting targeted peptide.

N.		Target Peptide Sequence Cysteine (Cys, C) : +57.0 Da (static modification) Glutamic acid (Glu, E) : +117.0 Da (variable modification)
1		ANTFLEEVR
2		ANTFLEEVR
3		ANTFLEEVR
4		ANTFLEEVR
5		ECVEETCSYEEAFEALESSTATDVFWAK
6		ECVEETCSYEEAFEALESSTATDVFWAK
7		ECVEETCSYEEAFEALESSTATDVFWAK
8		ECVEETCSYEEAFEALESSTATDVFWAK
9		ECVEETCSYEEAFEALESSTATDVFWAK
10		ECVEETCSYEEAFEALESSTATDVFWAK
11		ECVEETCSYEEAFEALESSTATDVFWAK
12		ECVEETCSYEEAFEALESSTATDVFWAK
13		ECVEETCSYEEAFEALESSTATDVFWAK
14		ECVEETCSYEEAFEALESSTATDVFWAK
15		ECVEETCSYEEAFEALESSTATDVFWAK
16		ECVEETCSYEEAFEALESSTATDVFWAK
17		ECVEETCSYEEAFEALESSTATDVFWAK
18		ECVEETCSYEEAFEALESSTATDVFWAK
19		ECVEETCSYEEAFEALESSTATDVFWAK
20		ECVEETCSYEEAFEALESSTATDVFWAK
21		ECVEETCSYEEAFEALESSTATDVFWAK
22		ECVEETCSYEEAFEALESSTATDVFWAK
23		ECVEETCSYEEAFEALESSTATDVFWAK
24		ECVEETCSYEEAFEALESSTATDVFWAK
25		ECVEETCSYEEAFEALESSTATDVFWAK
26		ECVEETCSYEEAFEALESSTATDVFWAK
27		ECVEETCSYEEAFEALESSTATDVFWAK
28		ECVEETCSYEEAFEALESSTATDVFWAK
29		ECVEETCSYEEAFEALESSTATDVFWAK
30		ECVEETCSYEEAFEALESSTATDVFWAK
31		ECVEETCSYEEAFEALESSTATDVFWAK
32		ECVEETCSYEEAFEALESSTATDVFWAK
33		ECVEETCSYEEAFEALESSTATDVFWAK
34	***in silico***	ECVEETCSYEEAFEALESSTATDVFWAK
35	**trypsin digestion**	ECVEETCSYEEAFEALESSTATDVFWAK
36		ECVEETCSYEEAFEALESSTATDVFWAK
37		ECVEETCSYEEAFEALESSTATDVFWAK
38		ECVEETCSYEEAFEALESSTATDVFWAK
39		ECVEETCSYEEAFEALESSTATDVFWAK
40		ECVEETCSYEEAFEALESSTATDVFWAK
41		ECVEETCSYEEAFEALESSTATDVFWAK
42		ECVEETCSYEEAFEALESSTATDVFWAK
43		ECVEETCSYEEAFEALESSTATDVFWAK
44		ECVEETCSYEEAFEALESSTATDVFWAK
45		ECVEETCSYEEAFEALESSTATDVFWAK
46		ECVEETCSYEEAFEALESSTATDVFWAK
47		ECVEETCSYEEAFEALESSTATDVFWAK
48		ECVEETCSYEEAFEALESSTATDVFWAK
49		ECVEETCSYEEAFEALESSTATDVFWAK
50		ECVEETCSYEEAFEALESSTATDVFWAK
51		ECVEETCSYEEAFEALESSTATDVFWAK
52		ECVEETCSYEEAFEALESSTATDVFWAK
53		ECVEETCSYEEAFEALESSTATDVFWAK
54		ECVEETCSYEEAFEALESSTATDVFWAK
55		ECVEETCSYEEAFEALESSTATDVFWAK
56		ECVEETCSYEEAFEALESSTATDVFWAK
57		ECVEETCSYEEAFEALESSTATDVFWAK
58		ECVEETCSYEEAFEALESSTATDVFWAK
59		ECVEETCSYEEAFEALESSTATDVFWAK
60		ECVEETCSYE[EAFEALESSTATDVFWAK
61		ECVEETCSYEEAFEALESSTATDVFWAK
62		ECVEETCSYEEAFEALESSTATDVFWAK
63		ECVEETCSYEEAFEALESSTATDVFWAK
64		ECVEETCSYEEAFEALESSTATDVFWAK
65		ECVEETCSYEEAFEALESSTATDVFWAK
66		ECVEETCSYEEAFEALESSTATDVFWAK
67		ECVEETCSYEEAFEALESSTATDVFWAK
68		ECVEETCSYEEAFEALESSTATDVFWAK
1		EEVRKGNL
2		EEVRKGNL
3		EEVRKGNL
4		EEVRKGNL
5		**ERECVEETCSY**
6		ERECVEETCSY
7		ERECVEETCSY
8		ERECVEETCSY
9		ERECVEETCSY
10	***in silico***	ERECVEETCSY
11	**chymotrypsin digestion**	ERECVEETCSY
12		ERECVEETCSY
13		ERECVEETCSY
14		ERECVEETCSY
15		ERECVEETCSY
16		ERECVEETCSY
17		ERECVEETCSY
18		ERECVEETCSY
19		ERECVEETCSY
20		ESSTATDVF
21		ESSTATDVF

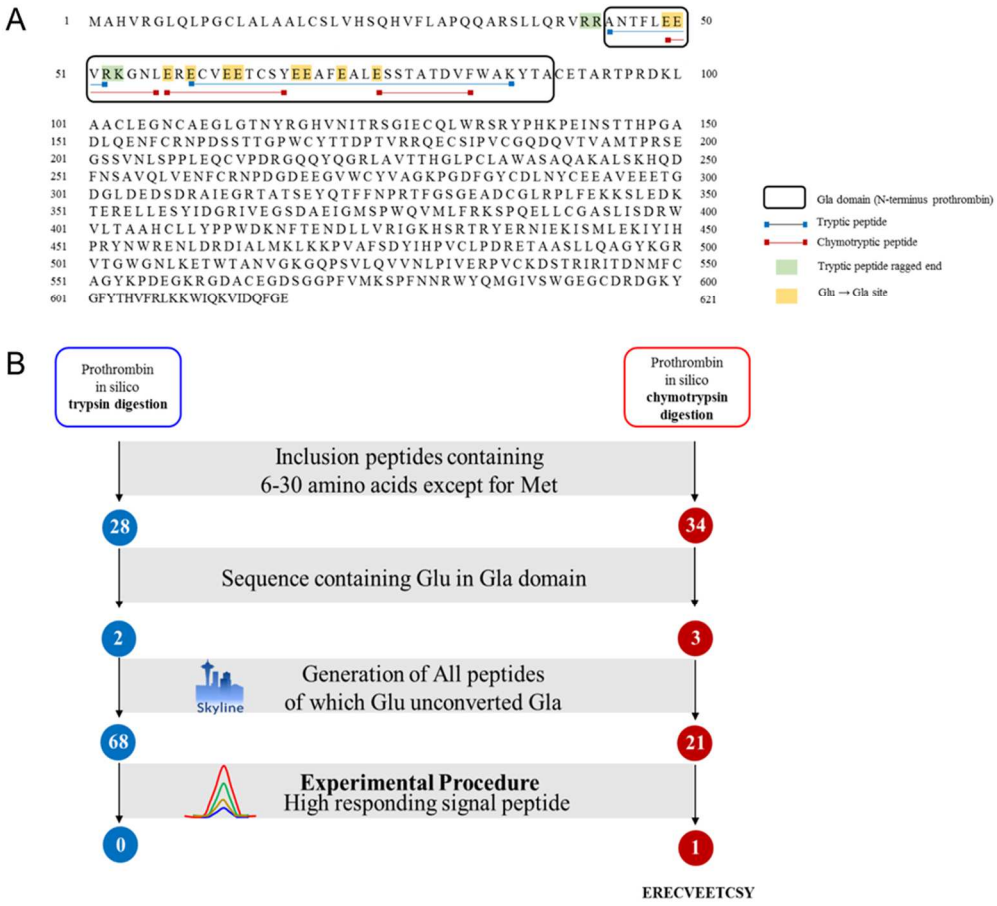

Figure 3. Scheme for determining the surrogate peptide for quantitating PIVKA-II

The Gla domain is located in the N-terminal region of prothrombin (A). In the Gla domain, containing 10 glutamate residues (Glu, E), a varying number of Glu residues are converted into Gla in patients with HCC. Two tryptic peptides (blue) and 3 chymotryptic peptides (red) were examined by MRM-MS analysis (B One chymotryptic peptide was selected to represent the concentration of PIVKA-II.

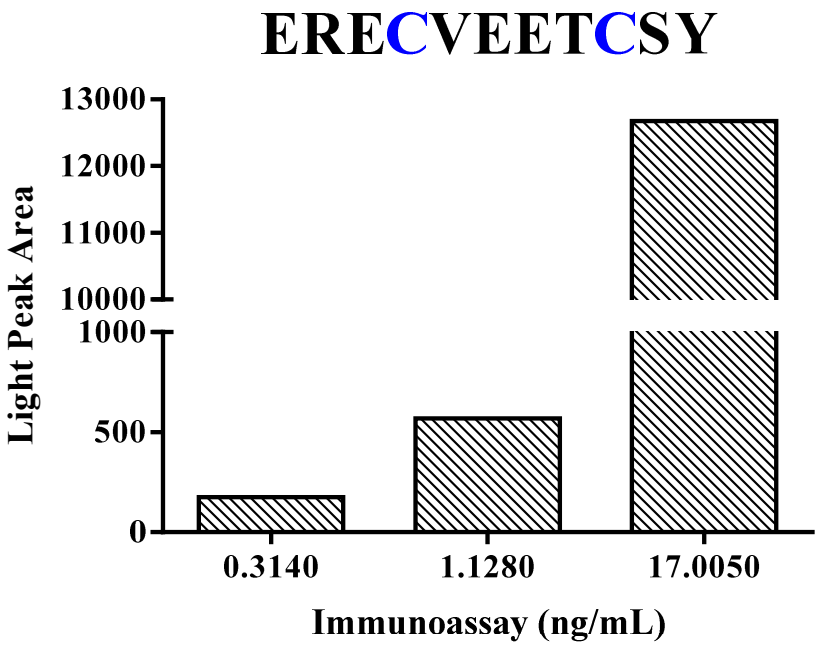

Figure 4. Intensity of the surrogate peptide from 3 patents with low, medium, high level of PIVKA-II

MRM-MS analysis result of ERECVEETCSY using 3 HCC clinical samples. Among 68 tryptic peptides and 21 chymotryptic peptides, a single peptide, ERECVEETCSY, was correlated with the result from immunoassay. X-axis values were concentrations of PIVKA-II from immunoassay and y-axis values were endogenous peak area from MRM-MS assay

Table 4. MRM-MS transitions for method development and quantification.

Analyte	peptide	Isotope type	Precursor ion (m/z)	Precursor ion charge state	Product ion (m/z)	Product ion type	Product ion charge state
AFP	GYQELLEK	unlabeled	490.3	2	759.4	**y6[b]**	1
					631.4	y5	1
					276.2	y2	1
					147.1	y1	1
					380.2	y6	2
					221.1	b2[c]	1
	GYQELLEK	labeled	494.3	2	767.4	**y6[b]**	1
					639.4	y5	1
					284.5	y2	1
					155.1	y1	1
					384.2	y6	2
					221.1	b2[c]	1
AFP-L3	V*D*[a]FTEIQK	unlabeled	490.3	2	880.4	y7	1
					765.4	**y6[b]**	1
					618.3	y5[c]	1
					517.3	y4	1
					383.2	y6	2
					215.1	b2	1
	V*D*[a]FTEIQK	labeled	494.3	2	888.5	y7	1
					773.4	**y6[b]**	1
					626.4	y5[c]	1
					525.3	y4	1
					387.2	y6	2
					215.1	b2	1
PIVKA-II	ERECVEETCSY	unlabeled	731.3	2	530.2	y4	1
					269.1	y2	1
					182.1	y1	1
					674.3	b5	1
					932.4	**b7[b]**	1
					466.7	b7	2
	ERECVEETCSY	labeled	736.3	2	540.2	y4	1
					279.1	y2	1
					192.1	y1	1
					674.3	b5	1
					932.4	**b7[b]**	1
					466.7	b7	2

[a] Asparagine (Asn, N), which was the original residue, was converted into aspartic acid (Asp, D) by PNGase-F.

[b] Product ions in bold are the best single transition chosen for quantification.

[c] Product ions are the second most intense transition chosen for qualification

3.1.3. Optimization of digestion and deglycosylation

To determine the optimal conditions for the trypsin digestion and PNGase-F deglycosylation, I examined the relationship between the digestion efficiency and the treated amount or the reaction time. Pooled HCC serum was immunoprecipitated with monoclonal AFP antibody and aliquoted.

Various amounts of trypsin (0, 1, 2, 4, 10, and 20 μg) were prepared using 6 aliquots, and the samples were digested at 37 ºC for 4 h. For various times (1, 2, 4, 8, 12, and 16 h), 6 aliquots were digested with 4 μg of trypsin at 37ºC. All digested samples were spiked with 100 fmol of stabled isotope-labeled synthetic peptide [GYQELLEK, obtained from 21st Century Biochemicals, Inc. (Marlborough, England), purity > 95% by AAA] as an internal standard. As shown in Figure 5 the peak area ratio was plotted versus the amount of trypsin and digestion time, which demonstrated that the efficiency of the digestion plateaued. Digestion was complete with 4 μg of trypsin and 4 h of digestion time, and these points were selected for the final optimal method.

Then, I evaluated the optimal amount of PNGase-F and the optimal reaction time for deglycosylation using AFP-enriched and AFP-L3-fractionated pooled HCC samples. Various amounts of PNGase-F (0, 1, 10, 100, 1000, and 5000 units) were prepared using 6 aliquots. The samples underwent deglycosylation at 37 ºC for 4 h. For various reaction times (1, 2, 4, 8, 16, and 24 h), 6 aliquots were deglycosylated with 1000 units of PNGase-F at 37ºC. Then, 12 samples were digested under optimal conditions (4 μg of trypsin and 4 h of digestion time) and spiked with 100 fmol of stable isotope-labeled synthetic peptide (VNFTEIQK) as an internal standard. As

shown in Figure 6 the peak area ratio was plotted versus the amount of PNGase-F and the reaction time, respectively, plateauing in deglycosylation efficiency. Deglycosylation was complete with 1000 units of PNGase-F and 4 h of deglycosylation time. These conditions were selected as the optimal method.

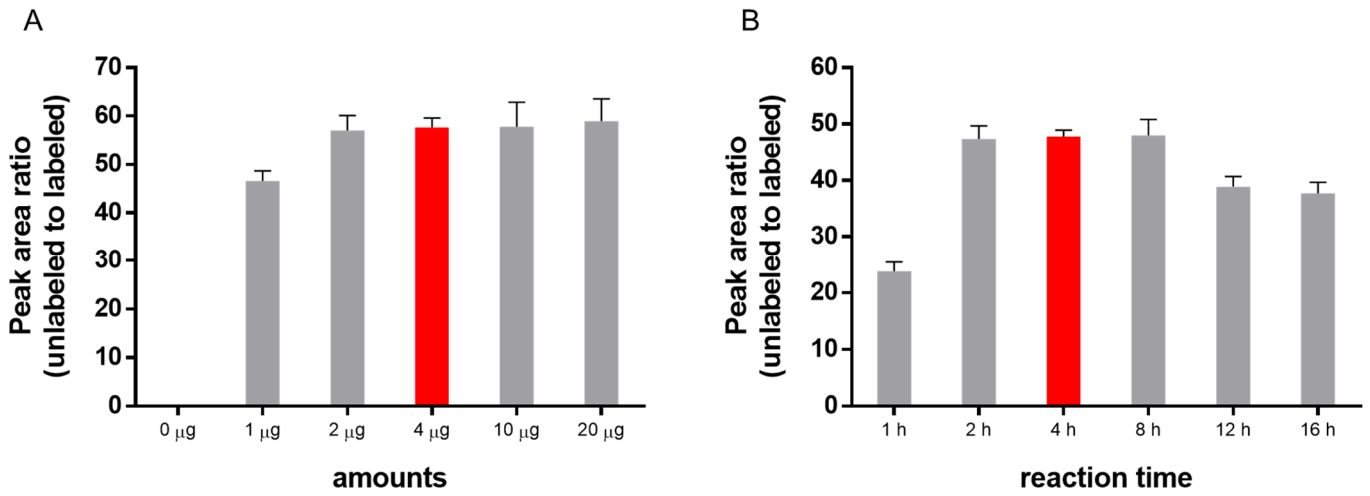

Figure 5. Optimization of digestion with trypsin.

To optimize the digestion with trypsin, various conditions were tested to determine the proper trypsin amount and reaction time. Different amounts (0, 1, 2, 4, 10, and 20 μg) of trypsin were tested, and the reaction time (4 h) was fixed (A). With regard to reaction time, many reaction times (1, 2, 4, 8, 12, and 16 h) were tested with a fixed amount of trypsin (4 μg) (B). The immunoprecipitated pooled HCC samples were digested and analyzed in triplicate. Four micrograms of trypsin and 4 h of reaction time were selected as the optimal conditions; these conditions were optimized on consideration of time and cost. Error bar represents the standard deviation of the 3 measurements.

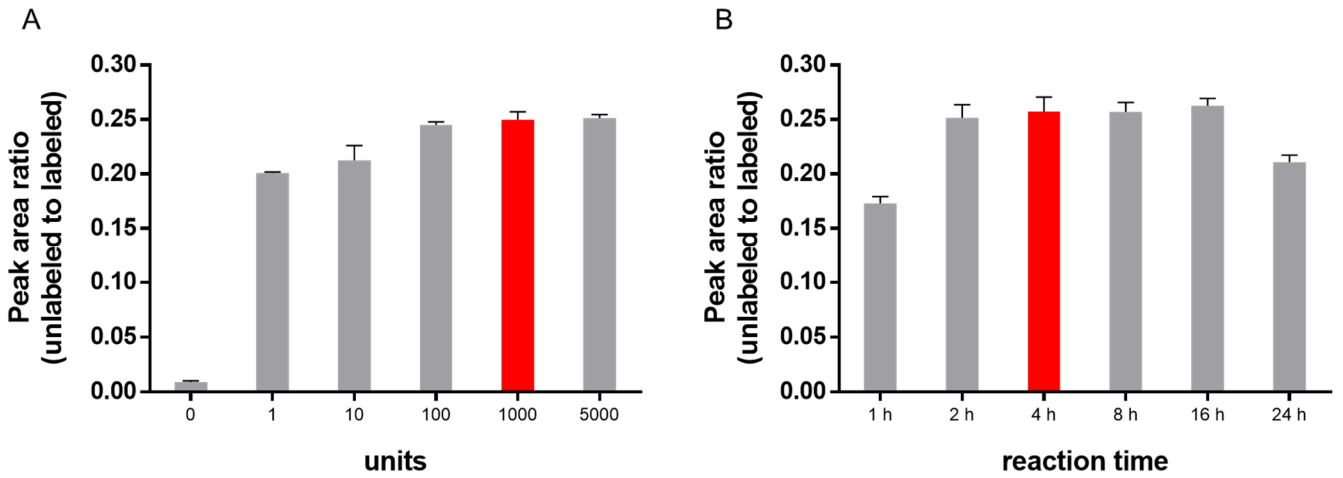

Figure 6. Optimization of deglycosylation with PNGase-F.

To optimize the deglycosylation with PNGase-F, various conditions were tested regard to the amount of PNGase-F and the reaction time. Several amounts (0, 1, 10, 100, 1000, and 5000 units) of PNGase-F were tested, with a fixed reaction time (4 h) (A). Many reaction times (1, 2, 4, 8, 16, and 24 h) were tested with a fixed amount of PNGase-F (1000 units) (B). The AFP-enriched and AFP-L3-fractionated pooled HCC samples were deglycosylated, digested and analyzed in triplicate. One thousand units of PNGase-F and 4 h of reaction time were selected as the optimal conditions; these conditions were optimized on consideration of time and cost. Error bar represents the standard deviation of the 3 measurements.

3.1.4. Assessment of complete deglycosylation

I also confirmed the deglycosylation efficiency by determining whether the sugar structures of AFP-L3 were completely removed under the optimized deglycosylation conditions (1000 units of PNGase-F and 4 h of deglycosylation). The AFP-enriched and AFP-L3-fractionated pooled HCC samples were divided into 2 samples (100 μL each). One sample was treated with PNGase-F according to the optimized PNGase-F treatment conditions, and the other sample was untreated. The 2 samples were analyzed by LiBA.

Between untreated and treated PNGase-F samples, the concentrations of AFP did not differ, whereas AFP-L3 was undetectable in PNGase-F treated sample (Figure 7). These results demonstrate that nearly all sugar structures on AFP were detached under the optimal deglycosylation conditions. The deglycopeptide had a single sequence, VDFTEIQK, although the sugar structures of AFP-L3 were heterogeneous, supporting that VDFTEIQK peptide is a true proxy for all heterogeneous fucose-attached sugar structures in AFP-L3. The MS/MS spectra of peptides that were obtained from the product ion scan mode are shown in Figure 8.

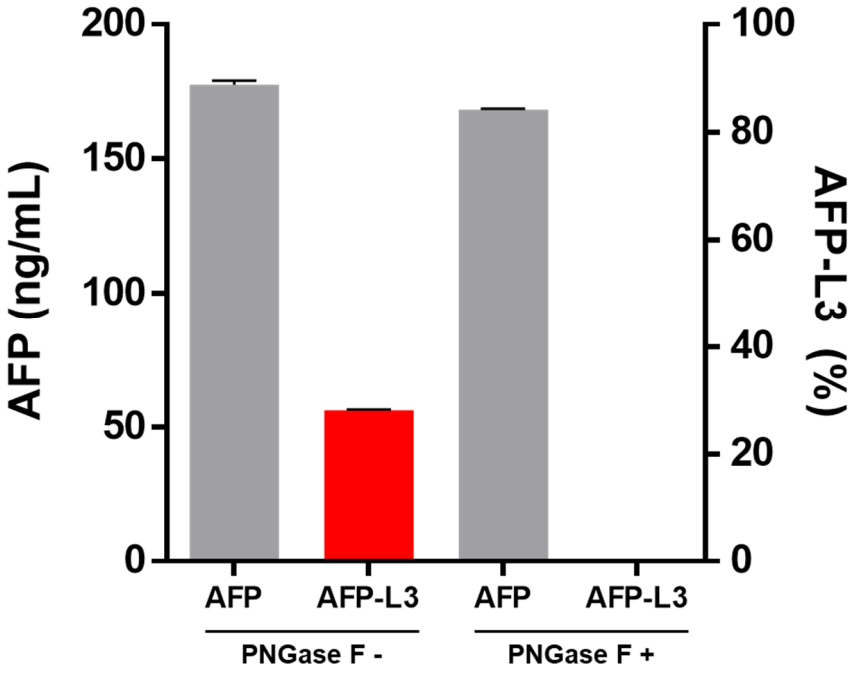

Figure 7. Verification of complete deglycosylation in the optimized deglycosylation method.

The optimized PNGase-F deglycosylation conditions were verified on another platform (liquid-phase binding assay, LiBA). PNGase-F-treated and untreated samples were compared with regard to the concentrations of AFP and AFP-L3. The AFP-enriched and AFP-L3-fractionated pooled HCC sample were analyzed in triplicate. AFP concentrations were similar between treatments, but AFP-L3 could not be detected after PNGase-F treatment. Thus, the heterogeneous fucose-attached sugar structures in AFP were all detached through the optimized deglycosylation method. Error bar represents the standard deviation of the 3 measurements.

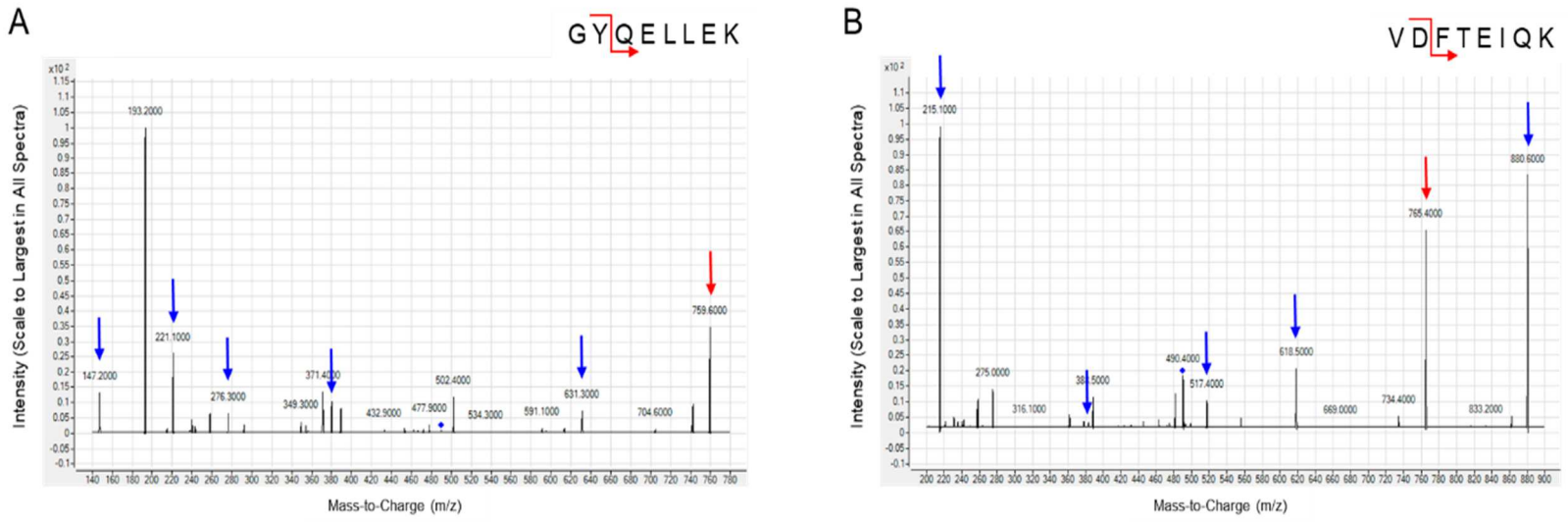

Figure 8. Product ion mass spectra of surrogate peptides.

Product ion scans for AFP and AFP-L3 surrogate peptides in HCC samples. Product ion mass spectra of endogenous GYQELLEK peptide (A), and VDFTEIQK peptide (B). Blue diamonds denote doubly charged precursor ions; red arrow signify singly charged quantification parent ion; blue arrows mark the 5 product ions used as qualification ions. The concentrations of AFP and AFP-L3 were 11.4 ng/mL and 5.7%, when measured by LiBA.

3.1.5. Diagnostic performance of the MRM-MS assay

As an HCC screening assay, the MRM-MS assay quantified the AFP concentration in all samples. However, LiBA failed to do so in 4 samples, because the AFP concentrations were below the LLOQ (400 cases versus 396 cases). The optimal cutoff value (the maximum sum of the diagnostic sensitivity and specificity) was calculated to best distinguish between the HCC and high-risk HCC group (chronic hepatitis plus liver cirrhosis). The optimal cutoff values for the AFP measurement were 6.00 ng/mL for the MRM-MS assay and 5.90 ng/mL for LiBA. The area under the receiver operating characteristic (AUROC) value, diagnostic sensitivity, and specificity were 0.766 (95% CI, 0.721–0.806), 76.0%, and 77.5% for the MRM-MS assay and 0.740 (95% CI, 0.694–0.782), 59.0%, and 89.5% for LiBA, respectively (Figure 9A-C and Table 5). With regard to AFP concentrations, the DeLong test failed to conclude that the MRM-MS assay differed significantly from LiBA (P = 0.070) in the HCC versus high-risk HCC group.

Of those patients with chronic hepatitis, AFP-L3 was detected by MRM-MS assay and LiBA in 3 and 1 patients, respectively, versus 18 and 19 patients with liver cirrhosis. In the HCC group, AFP-L3 was measured in 162 (81.0%) and 123 (61.5%) patients by MRM-MS assay and LiBA, respectively. For AFP-L3%, the optimal cutoff values of the MRM-MS assay and LiBA were 0.132% and 0.500% (each LLOQ concentrations), respectively. The AUROC, diagnostic sensitivity, and specificity were 0.854 (95% CI, 0.815–0.887), 81.0%, and 89.5% for the MRM-MS assay and 0.767 (95% CI, 0.722–0.807), 61.5%, and 90.0% for LiBA, respectively (Figure 9D-F and Table 5). The MRM-MS assay outperformed LiBA in quantifying

AFP-L3%, based on its significantly higher AUROC values in diagnosing HCC patients (DeLong test, $P < 0.0001$).

For AFP-L3%, 39 HCC samples were measured exclusively above the cutoff value (> 0.132%) only by MRM-MS assay, not by LiBA. Of the 39 HCC samples, 24 were below the cutoff value (≤ 5.90 ng/mL) when AFP was measured by LiBA, constituting false negative cases by LiBA. (Figure 10). Among 24 HCC samples, except in 1 patient, all AFP-L3% values in the 23 samples were ≤ 5.95% (range 0.41% to 5.95%; 22.1% for 1 patient), demonstrating that the MRM-MS assay readily identified small changes in AFP-L3 at low AFP concentrations due to the lower background noise and higher dynamic range in measuring AFP-L3 responses than with LiBA. The MRM-MS assay could determine a low cutoff value by quantifying low values of AFP-L3, resulting in a reduced false-negative rate, allowing effective HCC screening.

The AUROC values were 0.761 (95% CI, 0.707–0.814) and 0.721 (95% CI, 0.662–0.780) in distinguishing HCC from chronic hepatitis and liver cirrhosis, respectively, based on APIVKA-II measurements by SRM-MS assay, whereas the values by LiBA were 0.680 and 0.695, respectively (Figure 11). In all cases, the AUROC values by MRM-MS assay were higher (differing by 0.026–0.081) than those by LiBA. By DeLong test, the MRM-MS assay performed significantly better in distinguishing HCC from hepatitis than LiBA ($P = 0.0009$), whereas the assays performed similarly in differentiating HCC from liver cirrhosis ($P = 0.1361$). Overall, the MRM-MS assay outperformed LiBA in diagnosing HCC patients by quantifying PIVKA-II in serum samples.

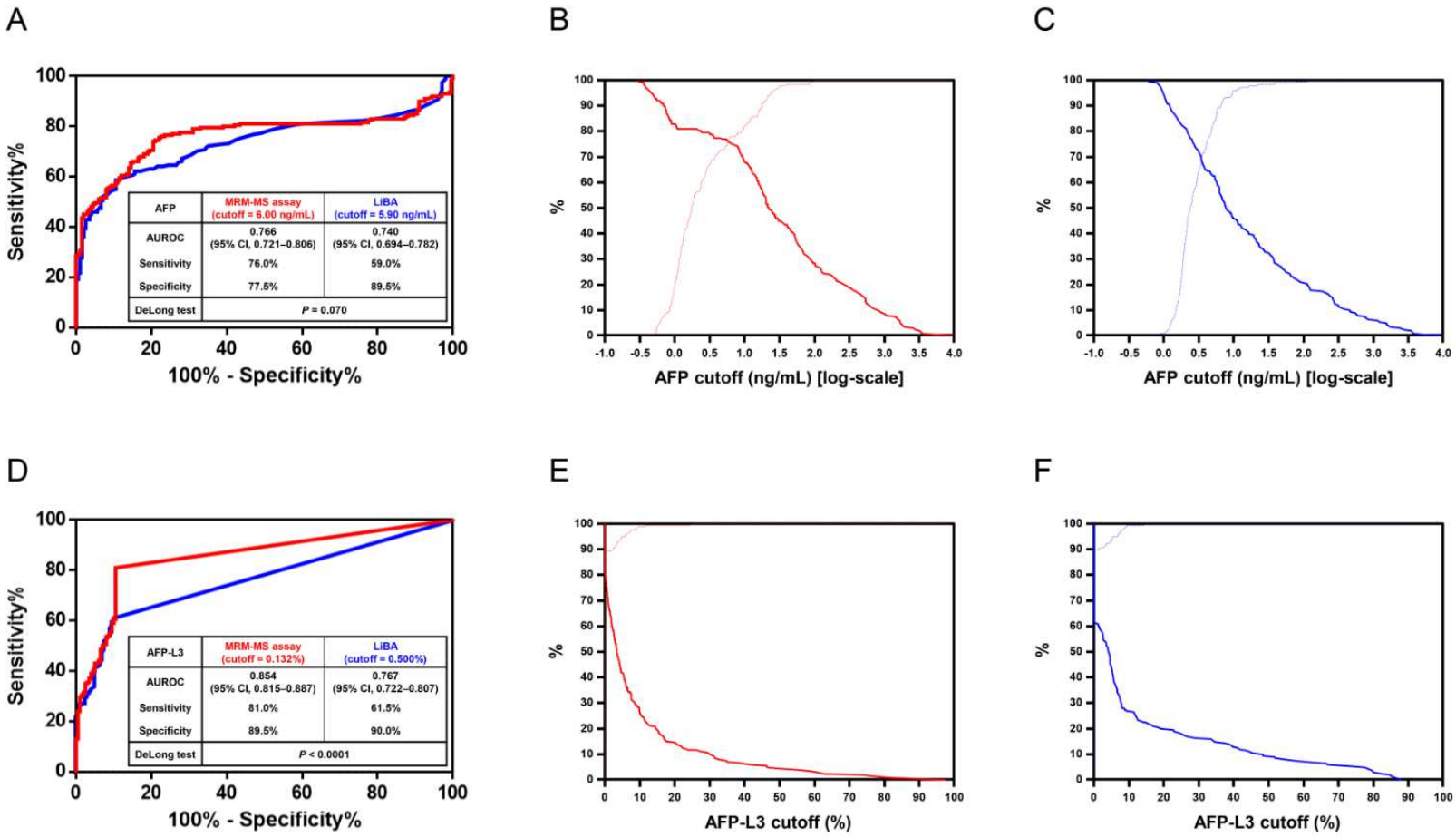
A
B
C
D
E
F
Sensitivity%
100% - Specificity%
AFP
MRM-MS assay (cutoff = 6.00 ng/mL)
LiBA (cutoff = 5.90 ng/mL)
AUROC
0.766 (95% CI, 0.721–0.806)
0.740 (95% CI, 0.694–0.782)
Sensitivity
76.0%
59.0%
Specificity
77.5%
89.5%
DeLong test
P = 0.070
%
AFP cutoff (ng/mL) [log-scale]
AFP-L3
MRM-MS assay (cutoff = 0.132%)
LiBA (cutoff = 0.500%)
0.854 (95% CI, 0.815–0.887)
0.767 (95% CI, 0.722–0.807)
81.0%
61.5%
89.5%
90.0%
P < 0.0001
AFP-L3 cutoff (%)

Figure 9. Comparison of AFP and AFP-L3 measurements between MRM-MS assay and LiBA.

For AFP (A–C) and AFP-L3 (D–F), ROC curves (A and D) and plot-versus-criterion values (B, C, E, and F) were plotted for HCC versus the high-risk HCC group (chronic hepatitis plus liver cirrhosis) by MRM-MS assay (red line) and LiBA (blue line), respectively. Delong test was used to determine whether the difference in area under the ROC values was significant. For the plot-versus-criterion values, the solid and dashed lines indicate the diagnostic sensitivity and specificity, respectively. The ROC curves for AFP-L3% formed a straight line due to the small number of detectable samples in the high-risk HCC group.

Table 5. Comparison of diagnostic power with cutoff value between assays.

Analyte	Assay type	AUROC	95% CI	Cutoff (ng/mL)	Analytical sensitivity	Analytical specificity
AFP						
	MRM-MS	0.766	0.721–0.806	<u>5.995</u>	76.00	77.50
				9.323	70.50	80.00
				17.83	56.50	90.00
				24.21	49.00	95.00
	LiBA	0.740	0.694–0.782	<u>5.900</u>	59.00	89.50
				4.750	63.50	80.00
				6.550	55.25	90.00
				9.567	46.50	95.00
AFP-L3						
	MRM-MS	0.854	0.815–0.887	<u>0.132</u>	81.00	89.50
				0.132	81.00	80.00
				2.101	60.50	90.00
				5.526	40.00	95.00
	LiBA	0.767	0.722–0.807	<u>0.500</u>	61.50	90.00
				0.500	61.50	80.00
				0.500	61.50	90.00
				6.700	34.50	95.00

AFP, alpha-fetoprotein; AFP-L3, lens culinaris agglutinin-reactive fraction of alpha-fetoprotein; MRM-MS, multiple reaction monitoring-mass spectrometry; LiBA, liquid-phase binding assay; AUROC, area under the receiver operating characteristic; CI, confidence interval.

Underline represents the optimal cutoff values.

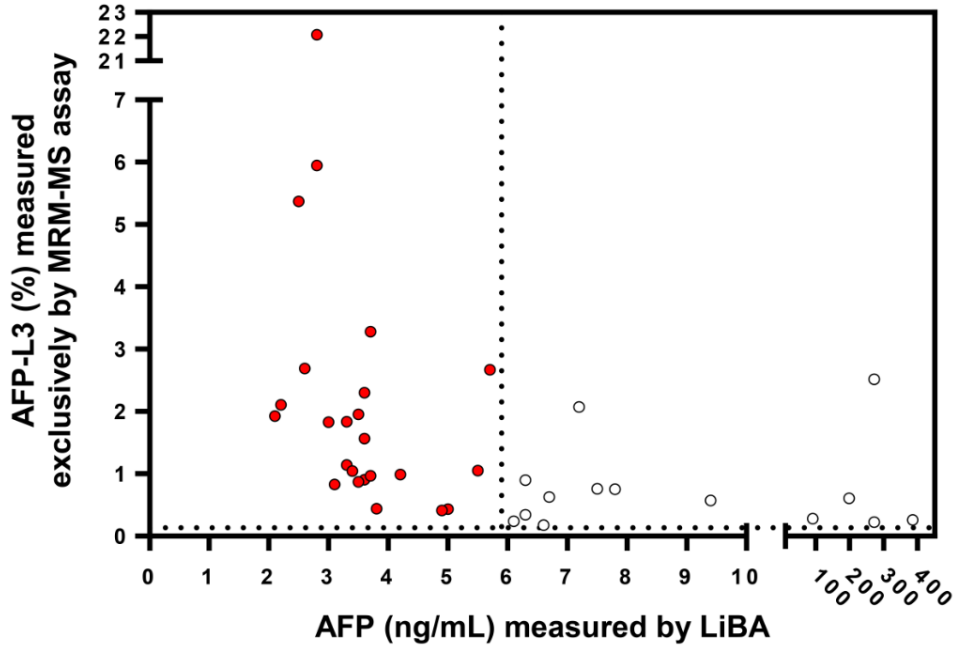

Figure 10. AFP-L3 measurements by MRM-MS assay compared to LiBA.

AFP-L3 values (y-axis) of the 39 HCC samples in which AFP-L3 was detected exclusively by MRM-MS assay but not LiBA, and AFP values (x-axis) as measured by LiBA for 39 HCC samples. The range of AFP-L3% was 0.18% to 22.1%, and the dotted line indicates the optimal cutoff values, which were 5.90 ng/mL and 0.132% for AFP by LiBA and AFP-L3 by MRM-MS assay, respectively. Further, 24 HCC samples (red dotted) were estimated to be below the cutoff of AFP by LiBA but were quantified to be above the cutoff of AFP-L3% by MRM-MS assay.

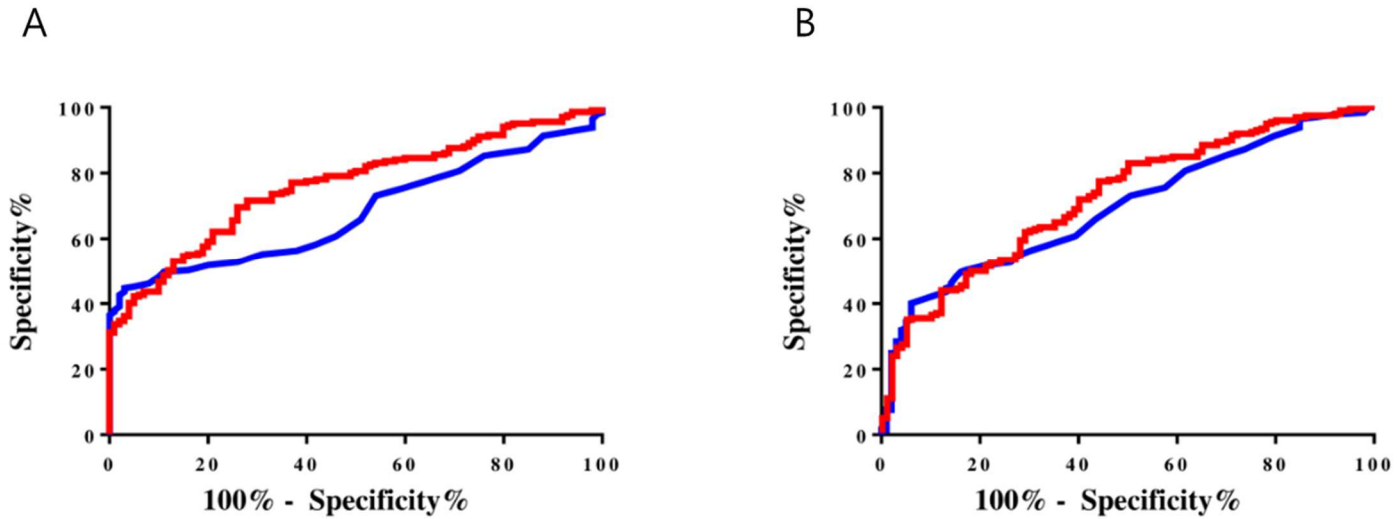

Figure 11. Comparison of PIVKA-II measurements between MRM-MS assay and LiBA.

For hepatitis versus HCC (A), the AUROC values were 0.680 and 0.761 for LiBA (blue) and SRM-MS (red), respectively. The optimal cutoff values were 24.50 ng/mL, 233.00 ng/mL for LiBA and MRM-MS, respectively, and the sensitivity and specificity were 51.78% and 80.00%, and 59.00% and 80.00%, respectively. For liver cirrhosis versus HCC (B), the AUROC values were 0.695 and 0.721 for LiBA and MRM-MS, respectively the optimal cutoff values were 24.50 ng/mL, 314.30 ng/mL for LiBA and MRM-MS, respectively, and the sensitivity and specificity were 51.78% and 78.79%, and 51.00% and 79.00%, respectively.

3.1.6. Method Comparison of the Two Assays

The performance of the SRM-MS assay was compared with that of the μTAS Wako i-30 autoanalyzer platform (Wako Pure Chemical Industries, Ltd, Osaka, Japan), which is the current FDA cleared standard liquid-phase binding assay (LiBA) system. The sample population, comprising chronic hepatitis (n=100), liver cirrhosis (n=100), and HCC (n=200), was analyzed by SRM-MS and LiBA. Because HCC is screened only in the high-risk HCC group, the reference interval for normal subjects was not evaluated. (Figure 12). The scatterplots and Deming regression equation for the comparison of AFP and AFP-L3 were LiBA = 0.799 [95% confidence interval (CI), 0.759–0.838] × MRM-MS + 0.286 for AFP when using 396 samples (Pearson correlation coefficient, R = 0.895; 95% CI, 0.873–0.913; P < 0.0001) and LiBA = 0.968 [95% CI, 0.851–1.084] x MRM-MS + 0.333 for AFP-L3 when using 141 samples (R = 0.813; 95% CI, 0.748–0.862; P < 0.0001), respectively (Figure 13 A–D and 14A–C and Table 6). In comparing the 2 assays using Bland-Altman plots, the MRM-MS assay had a mean positive bias of 0.340 (log2-scale in ng/mL; 95% CI, -2.569–3.249) for AFP and a mean negative bias of 0.231 (log2-scale in %; 95% CI, -2.200–1.738) for AFP-L3% (Figure 13E–H and 14D–F and Table 6). This result demonstrates that the LiBA underestimated the concentrations of AFP and AFP-L3 versus the MRM-MS assay.

The distribution of PIVKA-II concentrations in each group is shown in Figure 15. The 95% central ranges, as measured by SRM-MS and LiBA, were 21.3–494.0 ng/mL (median, 113.3 ng/mL) and 12.0–30.0 ng/mL (median, 19.0 ng/mL) for chronic hepatitis, respectively; 16.2–879.9 ng/mL (median, 111.1 ng/mL) and 1.8–

34.0 ng/mL (median, 7.8 ng/mL) for liver cirrhosis, respectively; and 44.3–149641.0 μg/mL (median 325.7 ng/mL) and 11.0–9708.0 ng/mL (26.0 ng/mL) for HCC, respectively. The range of the quantified distribution of SRM-MS was wider than that of LiBA. The expression of PIVKA-II in the HCC group, as measured by SRM-MS assay and LiBA, was significantly higher (Mann-Whitney test, $P < 0.001$) compared with the chronic hepatitis and liver cirrhosis groups. The Deming regression equation for the comparison of PIVKA-II was LiBA = 0.716 × SRM-MS - 0.226 (Pearson correlation coefficient, $R = 0.893$; 95% CI, 0.871–0.911; $P < 0.0001$), as in shown in Figure 16A. In comparing the 2 assays using Bland-Altman plots, the SRM-MS assay had a mean positive bias of 0.926 (log-scale in ng/mL, 95% CI, 0.010–1.843) for PIVKA-II (Figure 16B) .

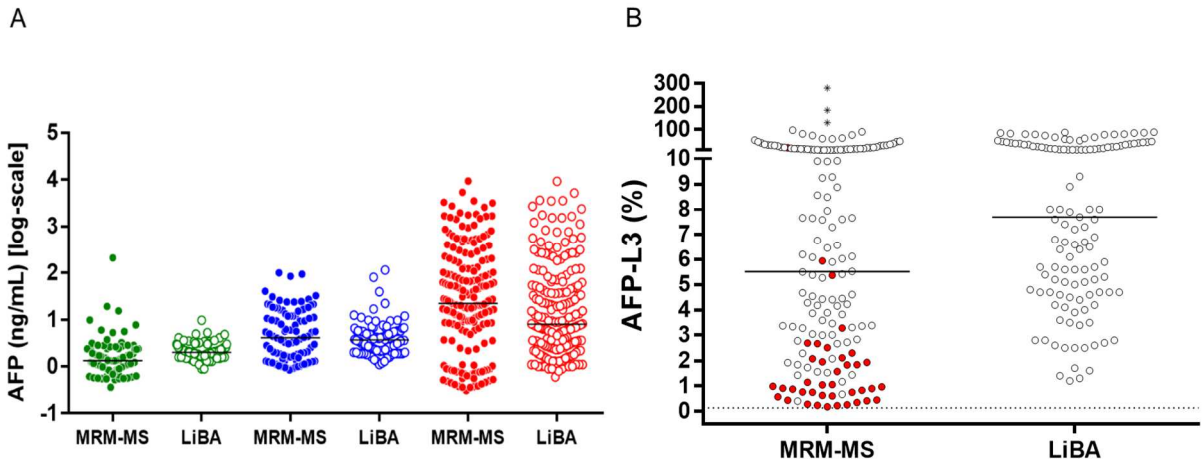

Figure 12. Quantitative distribution of AFP and AFP-L3(%)

Distribution of AFP concentrations for chronic hepatitis (n = 100, in green), liver cirrhosis (n = 100, in blue), and HCC (n = 200, in red) groups measured by the MRM-MS assay and LiBA (A). AFP-L3 was analyzed by two assays, and the comparison between MRM-MS assay and LiBA was conducted with regard to AFP-L3% in serum of HCC patients (B). AFP-L3% in the serum of HCC patients (red dot) was detected only by the MRM-MS assay. The middle line indicates the median value. The dashed line indicates the cutoff value (0.132%) of AFP-L3% by MRM-MS assay. The samples that were measured to be above 100% due to systemic errors are marked with an asterisk (*).

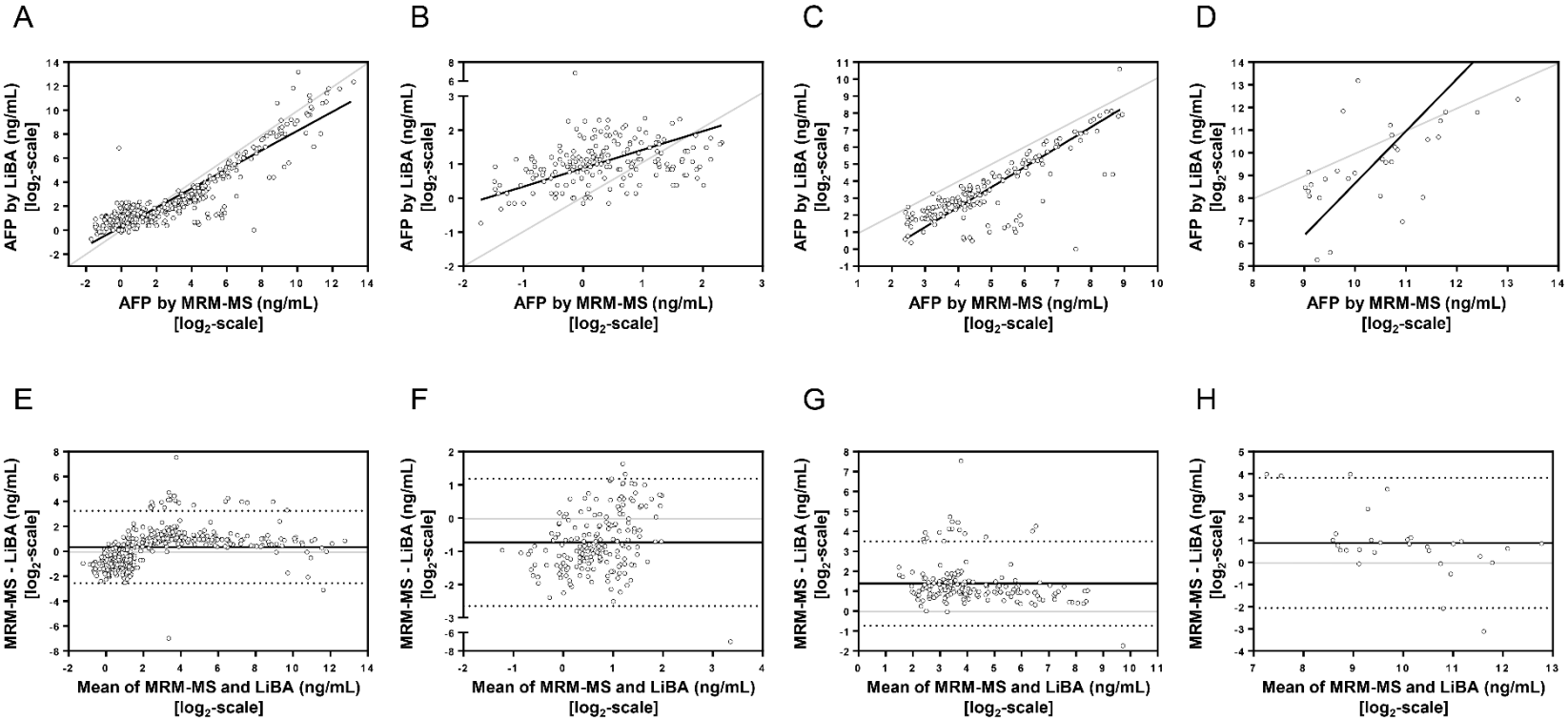
A
B
C
D
AFP by LiBA (ng/mL) [log2-scale]
AFP by MRM-MS (ng/mL) [log2-scale]
E
F
G
H
MRM-MS - LiBA (ng/mL) [log2-scale]
Mean of MRM-MS and LiBA (ng/mL) [log2-scale]

Figure 13. Correlation between AFP values as measured by MRM-MS assay and LiBA.

MRM-MS assay was compared with LiBA with regard to the concentration of AFP in 396 patients who were detectable by both assays. Deming regression (A–D) and Bland-Altman plots (E–H) were analyzed according to the concentration intervals of AFP, as measured by MRM-MS assay [entire concentration range (A, E); ≤ 5 ng/mL (B, F); 5–500 ng/mL (C, G); and > 500 ng/mL (D, H)]. The black line represents the mean difference between the 2 assays. The grey line represents unity (y=x), and the dotted line (dashed line) indicates the 95% CI of the mean difference.

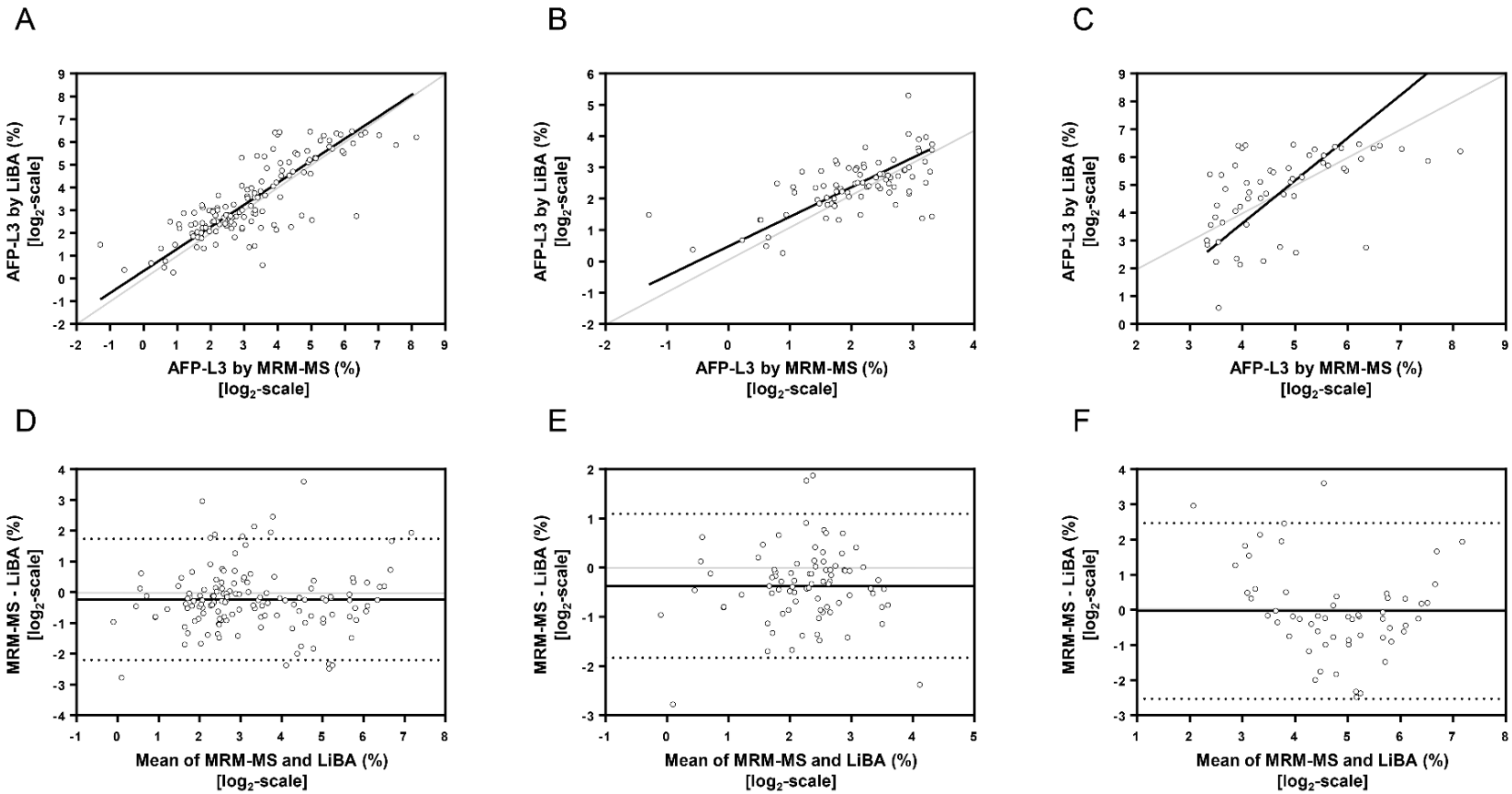
A
B
C
D
E
F
AFP-L3 by LiBA (%) [log2-scale]
AFP-L3 by MRM-MS (%) [log2-scale]
MRM-MS - LiBA (%) [log2-scale]
Mean of MRM-MS and LiBA (%) [log2-scale]

Figure 14. Correlation between AFP-L3 values as measured by MRM-MS assay and LiBA.

MRM-MS assay was compared with LiBA with regard to the concentration of AFP-L3 in 141 patients with HCC who were detectable by both assays. Deming regression (A–C) and Bland-Altman plots (D–F) were analyzed according to the concentration intervals of AFP-L3, as measured by MRM-MS assay [entire concentration range (A, D); ≤ 10% (B, E); and > 10% (C, F)]. The black line represents the mean difference between the 2 assays. The grey line represents unity (y=x), and the dotted line (dashed line) indicates the 95% CI of the mean difference.

Table 6. Correlation between assays by concentration interval of AFP and AFP-L3(%).

Analyte	Concentration range	No. of samples	Regression equation		R (95% CI)	Mean bias (95% CI)
			Slope (95% CI)	Y-intercept (95% CI)		
AFP						
	AMR	396	0.799 (0.759 to 0.838)	0.286 (0.107 to 0.464)	0.895 (0.873 to 0.913)	0.340 (-2.569 to 3.249)
	≤ 5 ng/mL	188	0.544 (0.284 to 0.803)	0.883 (0.637 to 1.128)	0.290 (0.154 to 0.416)	-0.733 (-2.651 to 1.184)
	5–500 ng/mL	177	1.178 (1.055 to 1.301)	-2.239 (-2.864 to -1.614)	0.819 (0.763 to 0.862)	1.385 (-0.729 to 3.499)
	> 500 ng/mL	31	2.300 (1.107 to 3.492)	-14.34 (-26.76 to -1.920)	0.591 (0.299 to 0.782)	0.881 (-2.055 to 3.817)
AFP-L3						
	AMR	141	0.968 (0.851 to 1.084)	0.333 (-0.081 to 0.747)	0.813 (0.748 to 0.862)	-0.231 (-2.200 to 1.738)
	≤ 10%	84	0.945 (0.693 to 1.196)	0.483 (-0.081 to 1.047)	0.636 (0.488 to 0.748)	-0.369 (-1.832 to 1.094)
	> 10%	57	1.535 (0.873 to 2.196)	-2.516 (-5.751 to 0.718)	0.531 (0.314 to 0.696)	-0.028 (-2.528 to 2.471)

AFP, alpha-fetoprotein; AFP-L3, lens culinaris agglutinin-reactive fraction of alpha-fetoprotein; AMR, analytical measurement range; CI, confidence interval; R, Pearson correlation coefficient.

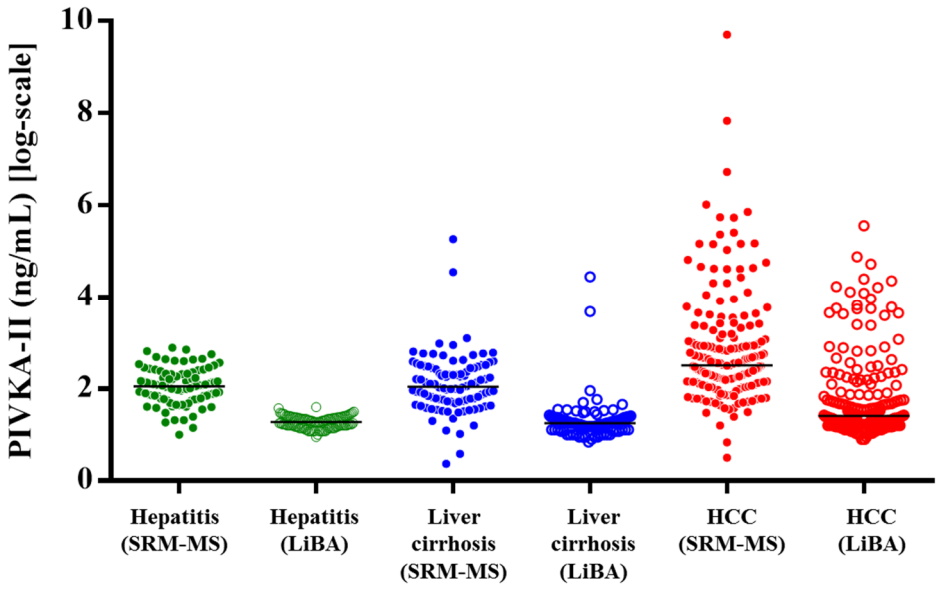

Figure 15. Quantitative distribution of PIVKA-II by MRM-MS assay and LiBA.

Distribution of PIVKA-II concentrations for chronic hepatitis (n=100, in green), liver cirrhosis (n=100, in blue), and HCC (n=200, in red) groups, as measured by MRM-MS and LiBA. Middle line indicates the median value.

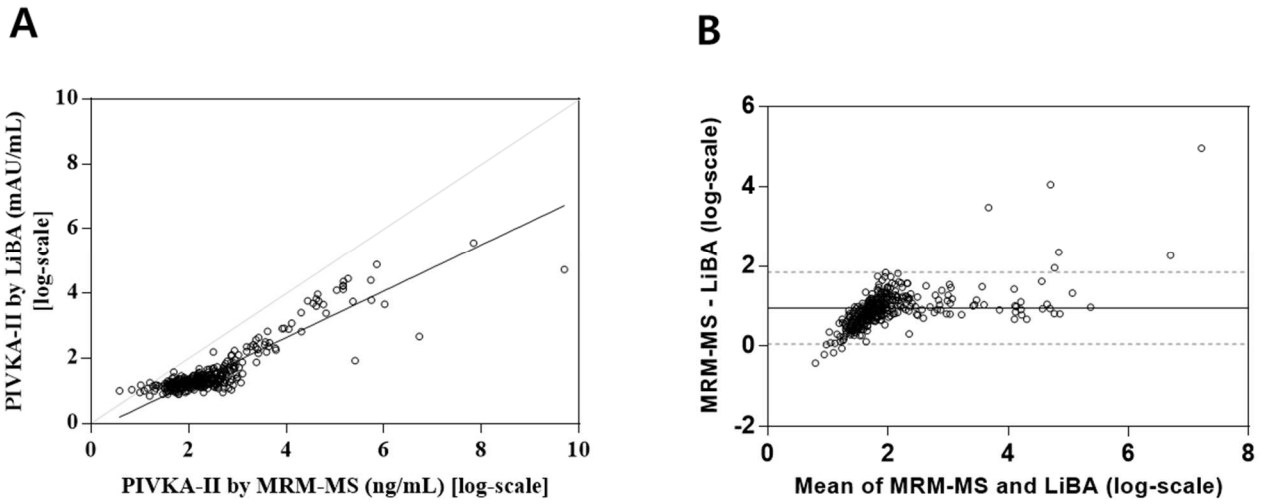

Figure 16. Correlation between PIVKA-II values as measured by SRM-MS assay and LiBA, respectively.

The gray line represents the line of identity (y=x), and the black line indicates the Deming line. The 95% confidence intervals (CIs) derived from the Deming slopes and intercepts are 0.680 to 0.752 and -0.320 to -0.131, respectively (A). Bland-Altman plot for PIVKA-II shows the difference in measured concentrations between SRM-MS and LiBA (B). The solid line represents the mean difference between assays, and the dotted line (dashed line) indicates the 1.96 standard deviation (SD) interval of the mean difference.

3.2. Method validation of the MRM-MS assay for AFP, AFP-L3 and PIVKA-II

3.2.1 Calibration curve

The limit of detection (LOD) and limit of quantification (LOQ) were determined, based on the mean values from the zero samples plus 3 and 10 times the standard deviation (SD), respectively. To determine the LLOQ, samples at 5 concentrations (GYQELLEK and VDFTEOIQK; 0.005–0.500 ng/mL, ERECVEETCSY; 0.10–5.00 ng/mL) that were close to the predetermined LOD were analyzed in 3 replicates in 1 d. The lowest concentration that met the precision [coefficient of variation (CV) < 20%], recovery of assay (within ± 15% of target), and signal-to-noise (S/N > 5) criteria was selected as the LLOQ. The LLOQs for GYQELLEK (representing AFP) and VDFTEIQK (representing AFP-L3) were both 0.051 ng/mL, and LLOQ for ERECVEETCSY (representing PIVKA-II) was 1.28 ng/mL.

A weighted 1/x linear regression model and logarithmic distribution weights were fitted to each dataset. The calibration curves showed good linearity, as evidenced by the coefficient of determination (R2) > 0.98 for AFP, AFP-L3 and > 0.99 for PIVKA-II (Figure 17). For each respective matrix, the bias from the mean slope were within ± 4% (AFP and AFP-L3), and ± 6% (PIVKA-II). In the reverse calibration curves, for the GYQELLEK and VDFTEIQK peptides, 7 of 8 calibrators (87.5%) had an acceptable bias within 15%, whereas, for ERECVEETCSY, all calibrators had an acceptable bias within 15%. The detailed results of the calibration curves are shown in Table 7, 8 and 9.

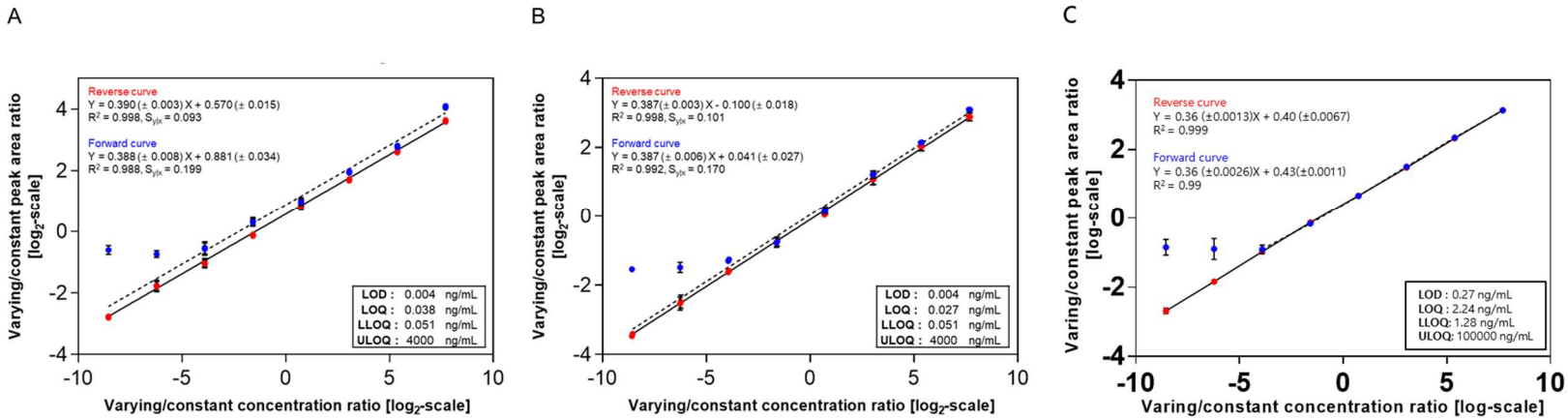

Figure 17 Evaluation of calibration curves.

The reverse and forward calibration curves for GYQELLEK (A), VDFTEIQK (B), ERECVEETCSY (C) were constructed with 8 concentration points using 6 matrices. The x-axis is the concentration ratio, transformed on a log2 scale, and the y-axis is the peak area ratio, transformed on a log2 scale. Error bars represent the standard deviations of the 3 measurements.

R2, coefficient of determination; Sy|x, standard deviation of residual.

Table 7. Results of reverse and forward calibration curve of GYQELLEK, VDFTEIQK peptides.

Calibration curve	Peptide	Matrix	Measurements	Blank sample	Zero sample	Calibrator 1	Calibrator 2	Calibrator 3	Calibrator 4	Calibrator 5	Calibrator 6	Calibrator 7	Calibrator 8
Reverse calibration curve	GYQELLEK		Expected conc. (ng/mL)	N/A	N/A	0.051	0.256	1.280	6.400	32.000	160.000	800.000	4000.000
		Matrix 1	Mean Area ratio	0.004	0.034	0.148	0.352	0.459	0.926	1.919	3.371	6.314	13.197
			SD	0.002	0.015	0.012	0.024	0.038	0.031	0.110	0.098	0.370	0.684
			CV (%)	43.647	43.558	7.786	6.838	8.265	3.302	5.712	2.903	5.860	5.181
		Matrix 2	Mean Area ratio	0.016	0.044	0.148	0.251	0.437	0.881	1.630	3.346	6.050	11.593
			SD	0.007	0.005	0.004	0.008	0.073	0.089	0.067	0.126	0.806	1.703
			CV (%)	40.648	11.261	2.736	3.039	16.758	10.088	4.107	3.751	13.317	14.691
		Matrix 3	Mean Area ratio	0.031	0.006	0.149	0.319	0.502	0.949	1.961	3.377	5.999	12.096
			SD	0.012	0.005	0.006	0.066	0.055	0.059	0.135	0.097	0.637	1.006
			CV (%)	37.632	80.972	3.836	20.772	10.925	6.220	6.893	2.878	10.624	8.320
		Matrix 4	Mean Area ratio	0.390	0.009	0.142	0.295	0.535	0.912	1.935	3.289	6.785	13.022
			SD	0.379	0.009	0.003	0.073	0.047	0.152	0.096	0.074	0.523	0.702
			CV (%)	97.374	97.748	1.874	24.882	8.751	16.684	4.986	2.259	7.701	5.387
		Matrix 5	Mean Area ratio	0.220	0.015	0.153	0.304	0.555	0.946	1.843	3.246	6.422	12.468
			SD	0.103	0.005	0.010	0.038	0.078	0.021	0.113	0.086	0.212	0.318
			CV (%)	46.791	35.223	6.546	12.535	13.975	2.235	6.125	2.649	3.306	2.549

	Matrix 6	Mean Area ratio	0.045	0.020	0.143	0.280	0.513	0.973	1.713	3.245	5.974	12.695
		SD	0.047	0.016	0.004	0.016	0.072	0.040	0.193	0.158	0.149	0.489
		CV (%)	104.094	83.351	2.800	5.664	14.053	4.122	11.263	4.866	2.494	3.850
	Mean[a]	Mean Area ratio	0.118	0.021	0.147	0.300	0.500	0.931	1.834	3.312	6.257	12.512
		SD	0.155	0.015	0.004	0.034	0.045	0.032	0.134	0.060	0.316	0.597
		CV (%)	131.988	69.386	2.937	11.474	8.974	3.452	7.310	1.823	5.045	4.772
	Bias[b]	Mean Area ratio	N/A	N/A	**100.405**	**128.293**	**94.524**	**91.749**	**103.750**	**94.637**	**97.010**	**114.621**
		SD	N/A	N/A	12.973	26.850	17.290	8.403	11.638	10.699	7.846	6.302
		CV (%)	N/A	N/A	**12.921**	**20.929**	**18.291**	**9.159**	**11.217**	**11.305**	**8.088**	**5.498**
		Signal-to-noise[c] (S/N)	N/A	N/A	**10.754**	**14.000**	**23.321**	**43.414**	**85.479**	**154.422**	**291.708**	**583.291**
V*D*FTEIQK												
	Matrix 1	Mean Area ratio	0.011	0.010	0.092	0.220	0.322	0.669	1.056	2.341	4.531	7.989
		SD	0.007	0.003	0.005	0.013	0.022	0.040	0.038	0.093	0.093	0.080
		CV (%)	61.198	33.544	5.073	5.781	6.943	5.980	3.581	3.953	2.060	1.002
	Matrix 2	Mean Area ratio	0.010	0.009	0.099	0.188	0.341	0.524	1.080	2.173	4.143	7.716
		SD	0.002	0.005	0.002	0.015	0.019	0.094	0.007	0.106	0.281	0.715
		CV (%)	18.565	50.229	1.838	8.232	5.578	18.026	0.621	4.865	6.772	9.263
	Matrix 3	Mean Area ratio	0.036	0.018	0.097	0.169	0.355	0.635	1.030	2.234	4.367	7.497
		SD	0.012	0.007	0.002	0.007	0.012	0.045	0.018	0.138	0.025	0.192
		CV (%)	33.992	42.659	2.561	4.414	3.285	7.103	1.755	6.164	0.563	2.562

		Matrix 4	Mean Area ratio	0.087	0.029	0.096	0.186	0.322	0.609	1.063	2.391	4.322	7.818
			SD	0.037	0.009	0.006	0.005	0.014	0.052	0.021	0.129	0.110	0.245
			CV (%)	42.973	31.102	5.768	2.682	4.188	8.541	1.959	5.386	2.551	3.132
		Matrix 5	Mean Area ratio	0.083	0.007	0.092	0.178	0.345	0.624	1.058	1.871	3.701	6.767
			SD	0.008	0.001	0.006	0.003	0.014	0.005	0.021	0.018	0.158	0.371
			CV (%)	9.503	17.899	6.213	1.601	4.097	0.807	1.956	0.984	4.261	5.480
		Matrix 6	Mean Area ratio	0.039	0.017	0.090	0.149	0.324	0.561	1.019	1.877	3.681	6.604
			SD	0.020	0.002	0.002	0.031	0.012	0.053	0.042	0.167	0.064	0.213
			CV (%)	52.626	12.493	2.255	20.776	3.782	9.371	4.150	8.886	1.730	3.218
		Mean[a]	Mean Area ratio	0.044	0.015	0.094	0.182	0.335	0.604	1.051	2.148	4.124	7.398
			SD	0.034	0.008	0.003	0.023	0.014	0.053	0.023	0.226	0.358	0.577
			CV (%)	76.307	55.361	3.663	12.888	4.155	8.748	2.147	10.503	8.668	7.798
		Bias[b]	Mean	N/A	N/A	**101.177**	**110.666**	**107.337**	**99.017**	**82.402**	**103.313**	**111.374**	**101.403**
			SD	N/A	N/A	11.331	25.904	15.894	16.142	10.884	12.469	5.376	6.445
			CV (%)	N/A	N/A	**11.200**	**23.407**	**14.807**	**16.302**	**13.208**	**12.070**	**4.827**	**6.356**
			Signal to noise[c] (S/N)	N/A	N/A	**8.161**	**12.384**	**22.827**	**41.141**	**71.650**	**146.411**	**281.149**	**504.361**
Forward calibration curve	GYQELLEK												
		Matrix 1	Mean Area ratio	24.636	0.557	0.610	0.599	0.672	1.201	1.811	3.665	7.368	17.228
			SD	20.154	0.027	0.070	0.057	0.044	0.130	0.163	0.457	0.436	1.401
			CV (%)	81.809	4.804	11.448	9.522	6.513	10.846	8.976	12.482	5.915	8.135

	Matrix 2	Mean Area ratio	144.076	0.549	0.688	0.672	0.755	1.436	1.955	3.917	6.580	17.640
		SD	113.589	0.163	0.094	0.103	0.082	0.058	0.111	0.247	0.084	0.375
		CV (%)	78.839	29.729	13.726	15.271	10.807	4.039	5.684	6.311	1.269	2.124
	Matrix 3	Mean Area ratio	0.536	0.603	0.776	0.660	0.628	1.192	2.059	4.363	7.325	15.920
		SD	0.078	0.092	0.022	0.071	0.189	0.113	0.442	0.399	0.613	0.338
		CV (%)	14.483	15.229	2.772	10.728	30.071	9.446	21.465	9.152	8.371	2.125
	Matrix 4	Mean Area ratio	49.804	0.508	0.684	0.562	0.745	1.242	2.219	4.023	6.997	17.453
		SD	67.515	0.039	0.022	0.028	0.086	0.240	0.108	0.301	0.231	0.397
		CV (%)	135.560	7.663	3.194	4.933	11.570	19.288	4.854	7.486	3.298	2.276
	Matrix 5	Mean Area ratio	103.628	0.541	0.603	0.657	0.826	1.179	1.985	3.850	7.079	17.120
		SD	68.153	0.035	0.014	0.060	0.068	0.019	0.187	0.149	0.534	0.390
		CV (%)	65.767	6.558	2.403	9.158	8.189	1.605	9.396	3.876	7.544	2.281
	Matrix 6	Mean Area ratio	59.813	0.634	0.724	0.587	0.590	1.373	2.173	3.943	6.925	17.580
		SD	32.084	0.101	0.080	0.119	0.134	0.064	0.204	0.445	0.692	1.000
		CV (%)	53.640	15.954	11.002	20.360	22.677	4.680	9.403	11.273	9.988	5.688
	Mean[a]	Mean Area ratio	63.749	0.565	0.681	0.623	0.703	1.270	2.034	3.960	7.046	17.157
		SD	52.494	0.046	0.067	0.046	0.088	0.108	0.150	0.231	0.289	0.638
		CV (%)	82.346	8.055	9.775	7.338	12.529	8.466	7.377	5.841	4.098	3.720
V*D*FTEIQK												
	Matrix 1	Mean Area ratio	29.302	0.404	0.370	0.357	0.408	0.587	1.149	2.307	4.535	8.567
		SD	17.387	0.026	0.076	0.035	0.036	0.037	0.115	0.136	0.138	0.357

	CV (%)	59.336	6.345	20.579	9.907	8.931	6.325	9.996	5.883	3.052	4.166
Matrix 2	Mean Area ratio	26.319	0.302	0.362	0.375	0.424	0.631	1.106	2.515	4.618	8.878
	SD	1.287	0.012	0.037	0.026	0.049	0.023	0.042	0.157	0.360	0.361
	CV (%)	4.892	3.858	10.296	7.011	11.476	3.703	3.773	6.243	7.804	4.069
Matrix 3	Mean Area ratio	0.416	0.362	0.345	0.313	0.409	0.584	1.260	2.272	4.307	8.449
	SD	0.048	0.009	0.044	0.016	0.050	0.083	0.065	0.031	0.272	0.189
	CV (%)	11.500	2.457	12.675	5.167	12.282	14.230	5.152	1.375	6.326	2.233
Matrix 4	Mean Area ratio	22.883	0.376	0.336	0.365	0.441	0.570	1.111	2.117	4.235	8.421
	SD	4.088	0.026	0.008	0.020	0.017	0.040	0.104	0.023	0.184	0.207
	CV (%)	17.866	6.957	2.475	5.544	3.801	6.932	9.355	1.078	4.348	2.453
Matrix 5	Mean Area ratio	32.121	0.303	0.350	0.388	0.419	0.628	1.044	2.575	4.523	8.300
	SD	9.641	0.024	0.034	0.043	0.014	0.033	0.038	0.161	0.078	0.330
	CV (%)	30.015	7.979	9.799	10.992	3.388	5.194	3.603	6.260	1.725	3.974
Matrix 6	Mean Area ratio	39.020	0.368	0.383	0.423	0.417	0.605	1.162	2.313	4.403	8.502
	SD	18.443	0.005	0.040	0.037	0.003	0.010	0.012	0.163	0.160	0.157
	CV (%)	47.264	1.317	10.427	8.631	0.797	1.674	0.991	7.063	3.642	1.850
Mean[a]	Mean Area ratio	25.010	0.353	0.358	0.370	0.420	0.601	1.139	2.350	4.437	8.520
	SD	13.239	0.041	0.017	0.036	0.012	0.025	0.072	0.168	0.147	0.197
	CV (%)	52.934	11.681	**4.872**	**9.863**	**2.884**	**4.128**	**6.348**	**7.164**	**3.317**	**2.311**

SD, standard deviation; CV, coefficient of variation; N/A, not applicable.

[a] The values were the mean of the peak area ratio from each of the 6 different matrices.

[b] Calculated by averaging the difference between nominal (expected) values and back-calculated values using each equation from the 6 matrices.

c Calculated by dividing the peak area ratio of each calculator (1–8) by that of the zero sample.

Underline indicates that the value did not meet the criteria.

Table 8. Results of reverse and forward calibration curve of ERECVEETCSY peptide.

Calibration curve	Matrix	Measurements	Blank sample	Zero sample	Calibrator 1	Calibrator 2	Calibrator 3	Calibrator 4	Calibrator 5	Calibrator 6	Calibrator 7	Calibrator 8
Reverse response curve		Expected conc. (ng/mL)	N/A	N/A	1.28	6.4	32	160	800	4000	20000	100000
	Matrix 1	Mean Area ratio	0.097	0.077	0.151	0.279	0.503	0.905	1.583	2.739	5.085	8.841
		STDEV	0.036	0.017	0.005	0.008	0.014	0.004	0.032	0.033	0.019	0.248
		CV (%)	37.689	21.962	3.196	2.727	2.844	0.479	2.015	1.22	0.366	2.81
	Matrix 2	Mean Area ratio	0.047	0.041	0.158	0.272	0.512	0.911	1.581	2.704	5.052	8.935
		STDEV	0.036	0.015	0.025	0.014	0.01	0.018	0.047	0.027	0.024	0.609
		CV (%)	76.142	35.899	15.848	5.159	1.906	1.949	2.97	0.996	0.467	6.813
	Matrix 3	Mean Area ratio	0.019	0.029	0.155	0.28	0.506	0.909	1.649	2.673	5.051	8.906
		STDEV	0.007	0.009	0.005	0.019	0.011	0.006	0.06	0.116	0.018	0.145
		CV (%)	37.774	30.534	3.548	6.732	2.145	0.614	3.639	4.346	0.35	1.629
	Matrix 4	Mean Area ratio	0.403	0.027	0.149	0.283	0.521	0.912	1.579	2.905	4.996	8.629
		STDEV	0.198	0.008	0.029	0.008	0.012	0.03	0.089	0.152	0.064	0.214
		CV (%)	49.001	30.087	19.678	2.71	2.215	3.255	5.612	5.227	1.286	2.474
	Matrix 5	Mean Area ratio	0.418	0.017	0.173	0.28	0.496	0.991	1.57	2.909	4.904	8.906
		STDEV	0.189	0.004	0.003	0.008	0.014	0.088	0.063	0.301	0.45	0.319
		CV (%)	45.138	26.294	1.988	2.706	2.898	8.908	4.034	10.363	9.185	3.587
	Matrix 6	Mean Area ratio	0.129	0.097	0.152	0.284	0.524	0.925	1.612	2.894	5.113	8.766

		STDEV	0.041	0.057	0.015	0.007	0.04	0.299	0.015	0.14	0.503	0.622
		CV (%)	31.57	58.726	9.828	2.321	7.698	32.344	0.936	4.83	9.839	7.091
	Average[a]	Mean Area ratio	0.185	0.048	0.156	0.28	0.51	0.925	1.596	2.804	5.034	8.831
		STDEV	0.179	0.032	0.009	0.004	0.011	0.033	0.03	0.11	0.074	0.116
		CV (%)	96.288	66.34	5.558	1.512	2.129	3.532	1.859	3.934	1.477	1.31
	Bias[b]	Mean Area ratio	N/A	N/A	**98.032**	**98.213**	**105.325**	**114.121**	**101.666**	**98.773**	**101.078**	**97.199**
		STDEV	N/A	N/A	8.362	6.485	9.289	10.663	6.117	10.102	3.492	5.058
		CV (%)	N/A	N/A	**8.53**	**6.603**	**8.82**	**9.343**	**6.017**	**10.228**	**3.454**	**5.204**
		Signal to noise[c] **(S/N)**	N/A	N/A	**5.001**	**6.944**	**12.661**	**22.957**	**39.581**	**69.555**	**124.863**	**219.052**
Forward response curve												
	Matrix 1	Mean Area ratio	3.016	0.604	0.538	0.742	0.513	0.909	1.697	2.85	5.079	8.914
		STDEV	2.058	0.077	0.148	0.188	0.203	0.057	0.149	0.314	0.594	0.574
		CV (%)	68.24	12.728	27.579	25.333	39.659	6.323	8.761	11.011	11.696	6.441
	Matrix 2	Mean Area ratio	6.336	1.319	0.737	0.566	0.519	0.89	1.531	2.855	4.914	8.571
		STDEV	5.002	0.631	0.115	0.05	0.065	0.151	0.217	0.274	1.464	1.434
		CV (%)	78.955	47.818	15.571	8.91	12.439	16.916	14.179	9.59	29.796	16.734
	Matrix 3	Mean Area ratio	0.61	10.395	0.575	0.578	0.509	0.899	1.529	2.847	5.073	9.004
		STDEV	0.113	7.715	0.047	0.278	0.08	0.23	0.292	0.292	0.96	0.543
		CV (%)	18.596	74.211	8.195	48.091	15.688	25.528	19.074	10.25	18.917	6.033
	Matrix 4	Mean Area ratio	18.839	0.427	0.55	0.445	0.516	0.9	1.588	2.817	5.123	8.712
		STDEV	3.828	0.067	0.119	0.075	0.043	0.163	0.3	0.025	0.767	1.598

	CV (%)	20.317	15.725	21.578	16.918	8.344	18.078	18.916	0.889	14.982	18.344
Matrix 5	Mean Area ratio	40.715	0.477	0.528	0.41	0.523	0.881	1.592	2.832	5.286	8.921
	STDEV	27.94	0.074	0.094	0.065	0.052	0.381	0.125	0.224	0.095	0.839
	CV (%)	68.623	15.51	17.73	15.83	9.924	43.174	7.88	7.909	1.789	9.4
Matrix 6	Mean Area ratio	20.005	0.771	0.455	0.551	0.636	0.892	1.559	2.901	5.185	8.702
	STDEV	1.814	0.317	0.068	0.098	0.206	0.156	0.188	0.179	0.399	1.097
	CV (%)	9.07	41.126	14.935	17.707	32.406	17.444	12.073	6.185	7.705	12.602
Average[a]	Mean Area ratio	14.92	2.332	0.564	0.549	0.536	0.895	1.583	2.851	5.11	8.804
	STDEV	15.012	3.963	0.094	0.117	0.049	0.009	0.062	0.028	0.125	0.167
	CV (%)	100.615	169.938	**16.631**	**21.351**	**9.188**	**1.051**	**3.92**	**0.997**	**2.438**	**1.896**

SD, standard deviation; CV, coefficient of variation; N/A, not applicable.

[a] The values were the mean of the peak area ratio from each of the 6 different matrices.

[b] Calculated by averaging the difference between nominal (expected) values and back-calculated values using each equation from the 6 matrices.

[c] Calculated by dividing the peak area ratio of each calculator (1–8) by that of the zero sample.

Underline indicates that the value did not meet the criteria.

Table 9. Comparison between matrices of the calibration curve.

Calibration curve	Peptide	Matrix	Slope		Bias (%) [a]	Intercept		$S_{y/x}$	R^2
			Mean	SD		Mean	SD		
Reverse calibration curve	GYQELLEK								
		Matrix 1	0.389	0.011	-0.223	0.611	0.060	0.169	0.995
		Matrix 2	0.393	0.006	0.848	0.468	0.031	0.089	0.999
		Matrix 3	0.384	0.006	-1.449	0.588	0.032	0.090	0.999
		Matrix 4	0.397	0.005	1.748	0.602	0.028	0.080	0.999
		Matrix 5	0.385	0.004	-1.238	0.605	0.023	0.064	0.999
		Matrix 6	0.391	0.004	0.239	0.547	0.024	0.067	0.999
		Mean	**0.390**	**0.005**		**0.570**	**0.055**	**0.093**	**0.998**
	VDFTEIQK								
		Matrix 1	0.391	0.010	0.969	-0.005	0.053	0.150	0.996
		Matrix 2	0.386	0.008	-0.134	-0.090	0.043	0.122	0.997
		Matrix 3	0.390	0.007	0.791	-0.066	0.035	0.099	0.998
		Matrix 4	0.393	0.006	1.574	-0.052	0.031	0.089	0.999
		Matrix 5	0.376	0.005	-2.823	-0.156	0.024	0.069	0.999
		Matrix 6	0.385	0.005	-0.498	-0.233	0.028	0.080	0.999
		Mean	**0.387**	**0.006**		**-0.100**	**0.082**	**0.101**	**0.998**
	ERECVEETCSY								
		Matrix 1	0.360	0.002	0.531	0.391	0.011	0.031	1.000
		Matrix 2	0.358	0.002	0.112	0.395	0.013	0.037	1.000
		Matrix 3	0.358	0.003	0.112	0.399	0.016	0.044	1.000
		Matrix 4	0.359	0.004	0.391	0.398	0.020	0.056	0.999
		Matrix 5	0.353	0.005	-1.481	0.428	0.024	0.068	0.999
		Matrix 6	0.359	0.003	0.363	0.410	0.015	0.042	1.000
		Average	**0.358**	**0.003**		**0.403**	**0.014**	**0.046**	**1.000**
Forward calibration curve	GYQELLEK								
		Matrix 1	0.397	0.022	2.199	0.810	0.097	0.213	0.988
		Matrix 2	0.374	0.027	-4.007	0.941	0.119	0.263	0.979
		Matrix 3	0.400	0.010	2.980	0.828	0.044	0.097	0.998
		Matrix 4	0.384	0.018	-1.242	0.930	0.081	0.178	0.991
		Matrix 5	0.377	0.025	-3.178	0.920	0.110	0.243	0.983
		Matrix 6	0.400	0.021	2.762	0.855	0.091	0.201	0.989
		Mean	**0.388**	**0.012**		**0.881**	**0.056**	**0.199**	**0.988**
	VDFTEIQK								
		Matrix 1	0.392	0.016	1.259	0.025	0.070	0.155	0.993
		Matrix 2	0.391	0.017	1.032	0.078	0.077	0.169	0.992
		Matrix 3	0.387	0.015	-0.095	0.036	0.065	0.143	0.994
		Matrix 4	0.380	0.022	-1.753	0.010	0.097	0.215	0.987
		Matrix 5	0.386	0.020	-0.147	0.054	0.088	0.194	0.989
		Matrix 6	0.386	0.015	-0.355	0.045	0.066	0.147	0.994
		Mean	**0.387**	**0.004**		**0.041**	**0.024**	**0.170**	**0.992**
	ERECVEETCSY								
		Matrix 1	0.354	0.004	0.007	0.440	0.016	0.035	1.000
		Matrix 2	0.351	0.003	-0.003	0.402	0.015	0.033	1.000
		Matrix 3	0.358	0.003	0.018	0.405	0.014	0.031	1.000
		Matrix 4	0.354	0.002	0.005	0.418	0.010	0.022	1.000
		Matrix 5	0.357	0.004	0.015	0.424	0.019	0.042	0.999
		Matrix 6	0.337	0.013	-0.043	0.503	0.055	0.122	0.995
		Average	**0.352**	**0.008**		**0.432**	**0.038**	**0.047**	**0.999**

SD, standard deviation; $S_{y/x}$, standard deviation of residual; R^2, coefficient of determination.

[a] Bias (%) = $\frac{\text{(slope of each of matrix - mean slope)}}{\text{mean slope}} \times 100$

3.2.1.1. Estimation of the LLOQ of AFP-L3% by MRM-MS assay

To estimate the LLOQ of AFP-L3%, I mixed the 2 HCC serum samples, because there was a lack of authentic standard material for the glycosylated form of AFP-L3. Serum A (high AFP-L3) and B (low AFP-L3) were selected, based on the concentrations as measured by LiBA. Two HCC serum samples were mixed at varying proportions (100:0, 0:100, 1:99, 2:98, 3:97, 4:96, and 5:95). All samples were analyzed in 3 replicates in 1 d (Table 10)

The measured concentration of the unmixed samples and the ratio at which they were mixed were taken into account when calculating the expected concentration of the mixed sample. The bias between the expected and measured values ranged from 0.132% to 0.222% for AFP-L3, which was within the satisfactory range of ± 15%. The lowest concentration (0.132%) with low bias and high precision (CV < 15%) was considered to be the LLOQ of AFP-L3% by MRM-MS assay.

Table 10. Lower limit of quantification evaluation of AFP-L3% in clinical samples.

Mixing ratio	Sample A Sample B	100.0 0.0	0.0 100.0	1.0 99.0	2.0 98.0	3.0 97.0	4.0 96.0	5.0 95.0
AFP (ng/mL)	Expected conc. (ng/mL)[a]	452.816	772.943	769.742	766.540	763.339	760.138	756.937
	Measured conc. (ng/mL)[b]	452.816	772.943	835.358	716.627	736.629	780.700	847.419
	SD	15.023	24.320	78.158	23.135	44.633	64.457	34.473
	CV (%)	3.318	3.146	9.356	3.228	6.059	8.256	4.068
	Bias (%)[c]	0.000	0.000	8.524	-6.512	-3.499	2.705	11.954
AFP-L3 (ng/mL)	Expected conc. (ng/mL)[a]	33.617	0.000	0.336	0.673	1.009	1.345	1.681
	Measured conc. (ng/mL)[b]	33.617	0.000	0.032	0.393	0.951	1.413	1.864
	SD	1.887	NA	0.019	0.068	0.070	0.097	0.072
	CV (%)	5.612	NA	59.477	17.399	7.342	6.888	3.859
	Bias (%)[c]	0.000	0.000	-90.616	-41.522	-5.714	5.094	10.861
AFP-L3 (%)	Expected conc. (%)[a]	7.422	0.000	0.044	0.088	**0.132**	0.177	0.222
	Measured conc. (%)[b]	7.422	0.000	0.004	0.055	0.129	0.181	0.220
	SD	0.260	NA	0.002	0.009	0.010	0.006	0.018
	CV (%)	3.508	NA	56.949	16.534	7.599	3.118	8.066
	Bias (%)[c]	0.000	0.000	-91.569	-37.482	-2.152	2.424	-0.759

standard deviation; CV, coefficient of variation; NA, not applicable.

[a] Calculated by measuring the concentration of 2 unmixed samples and then applying the mixed ratio of the 2 samples.

[b] Calculated based on mean value of 3 replicates.

[c] Bias (%) = $\frac{\text{(measured concentration - expected concentration)}}{\text{expected concentration}} \times 100$

3.2.2. Analytical specificity (selectivity or interference)

Analytical specificity was deemed to be satisfactory per FDA, KFDA, and EMA standards if the calculated percentages of the interferents were less than 20% of that of the LLOQ sample and if the peak areas of the internal standard were less than 5% of those of the LLOQ. The data that I used to calculate the actual analytical specificity can be found in Table 11.

The CLSI guidelines dictate that if the mean ion ratio [(qualifier ion signal (2nd most intense ion))/(quantifier ion signal (most intense ion))] of the internal standard is above 50% in all patient samples, the ion ratio for each patient sample should be within 20% of the mean ion ratio. For the GYQELLEK and VDFTEIQK peptides in all patient samples, the mean ion ratios were 70.7% and 97.8%, respectively, and 8 patients (2.0% for GYQELLEK peptide) and 6 patients (1.5% for VDFTEIQK peptide) had deviations in %difference of over 20% from the mean ion ratios.

All blank samples were examined for chromatographic interference at the expected retention times of the analyte and internal standard. Under the optimized MRM-MS conditions, no interference from the matrix was observed at the retention time for GYQELLEK and VDFTEIQK. In addition, if the peaks of interest become split in the LC chromatogram, the analyte or internal standard peptides are influenced by interfering substances. Of the 400 patient samples, split peaks were observed in the LC chromatogram for 0.50% (n = 2) and 14.75% (n = 59) of samples for the GYQELLEK and VDFTEIQK peptides, respectively.

For ERECVEETCSY peptide, the specificity of the b7 ion at a single charge did not meet the IS criteria (Table 12). The b7 ion at a single charge was selected to quantify a product ion, based on its highest intensity. In the method validation experiments, the results with the b7 product ion met all of the criteria except for the specificity of the internal standard. The tyrosine (Tyr, Y) residue in the labeled target peptide (ERECVEETCSY) was labeled. The b7 ions of the labeled and unlabeled peptide had the same m/z values, because the b7 ion did not contain the tyrosine at the C-terminus. Despite precursor ions having different m/z values and being selected separately, corresponding to each m/z value in the first quadrupole (Q1), the same m/z of the product ions can actually cause interference in Q3.

Thus, I evaluated the specificity using the 2 other product ions (the second- and third-most intense ions). The results are summarized in Table 12. Using the y4 (1+ ion charge state) and y2 (1+ ion charge state) product ions, the specificity for the analyte and internal standard met the criteria. The y4 and y2 ions had different m/z values in the labeled and unlabeled peptide, respectively, and the specificity% values met the criteria for the analyte and internal standard. Most of the results from the method validation met the criteria using the b7 ion, except for the specificity of the internal standard. Considering the overall validation of this quantitative clinical application, the quantification using the b7 ion was suitable.

Table 11. Analytical specificity (Interference) results of GYQELLEK, VDFTEIQK from blank sample analysis by MRM-MS assay.

Peptide	Matrix	Peptide type[a]	Peak area of blank sample[b]	Peak area of LLOQ	Interference (%)[c]
GYQELLEK		unlabeled			
	Matrix 1		241.667	1622.333	14.896
	Matrix 2		17.667	6156.667	0.287
	Matrix 3		17.667	1471.333	1.201
	Matrix 4		4.333	831.333	0.521
	Matrix 5		5.000	1248.667	0.400
	Matrix 6		37.000	1834.333	2.017
	Mean ± SD		**53.889 ± 92.753**	**2194.111 ± 1971.378**	**3.22 ± 5.757**
		labeled			
	Matrix 1		3.667	233.333	1.571
	Matrix 2		1.000	896.000	0.112
	Matrix 3		1.667	219.000	0.761
	Matrix 4		7.000	117.000	5.983
	Matrix 5		2.667	185.667	1.436
	Matrix 6		6.333	262.000	2.417
	Mean ± SD		**3.722 ± 2.462**	**318.833 ± 287.088**	**2.047 ± 2.080**
VDFTEIQK		unlabeled			
	Matrix 1		151.000	1624.667	9.294
	Matrix 2		91.000	2785.333	3.267
	Matrix 3		18.667	5438.000	0.343
	Matrix 4		43.000	2049.000	2.099
	Matrix 5		27.000	6989.667	0.386
	Matrix 6		56.000	4972.333	1.126
	Mean ± SD		**64.444 ± 49.449**	**3976.500 ± 2138.789**	**2.753 ± 3.392**
		labeled			
	Matrix 1		6.667	155.333	4.292
	Matrix 2		14.667	275.000	5.333
	Matrix 3		3.000	528.000	0.568
	Matrix 4		11.000	199.333	5.518
	Matrix 5		5.667	640.667	0.884
	Matrix 6		5.333	449.000	1.188
	Mean ± SD		**7.722 ± 4.297**	**374.556 ± 194.204**	**2.964 ± 2.329**

[a] Stable isotope-unlabeled peptide as internal standard and stable isotope-labeled peptide as target analyte.

[b] For blank sample, the peak area was from endogenous peptide or interferents corresponding to m/z of the designated peptides.

[c] Calculated by dividing the peak area of the blank sample by that of the LLOQ (calibrator 1) and multiplying the result by 100.

Table 12. Analytical specificity (Interference) results of ERECVEETCSY from blank sample analysis by MRM-MS assay.

Peptide	Matrix	Product ion	Peptide type[a]	Peak area of Blank sample[b]	Peak area of LLOQ	Interference (%)[c]
ERECVEETCSY		b7 (1+)	unlabeled			
	Matrix 1			127.667	857.333	14.891
	Matrix 2			165.667	267.333	61.970
	Matrix 3			230.667	1095.667	21.053
	Matrix 4			43.333	359.000	12.071
	Matrix 5			92.667	1263.333	7.335
	Matrix 6			121.000	1472.000	8.220
	Average					**20.923**
			labeled			
	Matrix 1			14.000	129.667	10.797
	Matrix 2			7.000	41.000	17.073
	Matrix 3			4.333	168.333	2.574
	Matrix 4			31.667	56.000	56.548
	Matrix 5			40.000	218.333	18.321
	Matrix 6			12.667	225.667	5.613
	Average					**18.488**
		y4 (1+)	unlabeled			
	Matrix 1			8.333	581.333	1.433
	Matrix 2			5.667	236.667	2.394
	Matrix 3			9.000	696.333	1.292
	Matrix 4			3.000	201.667	1.488
	Matrix 5			7.667	972.000	0.789
	Matrix 6			2.333	1146.333	0.204
	Average			**6.000**	**639.056**	**1.267**
			labeled			
	Matrix 1			5.000	51.667	9.677
	Matrix 2			1.000	69.000	1.449
	Matrix 3			0.667	131.667	0.506
	Matrix 4			6.000	56.000	10.714
	Matrix 5			10.333	134.333	7.692
	Matrix 6			2.667	159.333	1.674
	Average			**4.278**	**100.333**	**5.286**
		y2 (1+)	unlabeled			
	Matrix 1			25.667	1758.667	1.459
	Matrix 2			12.000	557.667	2.152

Matrix 3		6.333	2233.000	0.284
Matrix 4		9.000	947.333	0.950
Matrix 5		21.000	3771.000	0.557
Matrix 6		17.000	5013.667	0.339
Average		**15.167**	**2380.222**	**0.957**
	labeled			
Matrix 1		2.333	227.333	1.026
Matrix 2		4.333	211.000	2.054
Matrix 3		2.333	523.000	0.446
Matrix 4		7.000	171.333	4.086
Matrix 5		12.000	594.333	2.019
Matrix 6		5.333	725.667	0.735
Average		**5.556**	**408.778**	**1.728**

[a] Stable isotope-unlabeled peptide as internal standard and stable isotope-labeled peptide as target analyte.

[b] For blank sample, the peak area was from endogenous peptide or interferents corresponding to m/z of the designated peptides.

[c] Calculated by dividing the peak area of the blank sample by that of the LLOQ (calibrator 1) and multiplying the result by 100.

.

3.2.3. Analytical sensitivity

For the calculation of S/N, the signal and noise were represented by the peak area ratios of calibrator 1 and the zero sample (not blank sample) in each matrix. At the level of LLOQs of GYQELLEK and VDFTEIQK (both 0.051 ng/mL) and ERECVEETCSY (1.28 ng/mL) in serum, the signal responses were over 5-fold but less than 20 times the values in the zero sample. The precision and recovery of assay were within 20% for the 3 peptides (Table 13). The analytical sensitivity of my method for measuring AFP and AFP-L3 concentrations in the 6 serum samples met the requirements of the guidelines.

Table 13. Analytical sensitivity results from zero sample analysis with MRM-MS assay.

Peptide	Matrix	Measurements	Zero sample	LLOQ	Signal-to-noise (S/N)[a]	(%)[b]	CV (%)
GYQELLEK							
	Matrix 1	PAR	0.034	0.148	4.297		
		Conc. (ng/mL)	0.002	0.052		102.213	19.718
	Matrix 2	PAR	0.044	0.148	3.349		
		Conc. (ng/mL)	0.002	0.051		100.39	7.097
	Matrix 3	PAR	0.006	0.149	23.754		
		Conc. (ng/mL)	0	0.053		103.355	9.983
	Matrix 4	PAR	0.009	0.142	15.837		
		Conc. (ng/mL)	0	0.046		90.426	4.777
	Matrix 5	PAR	0.015	0.153	10.004		
		Conc. (ng/mL)	0	0.057		111.601	16.425
	Matrix 6	PAR	0.02	0.143	7.282		
		Conc. (ng/mL)	0.001	0.047		92.47	7.246

	Mean	PAR	0.021	0.147	**10.754**		
		Conc. (ng/mL)	0.001	0.051		**100.076**	**10.874**
V*D*FTEIQK							
	Matrix 1	PAR	0.01	0.092	9.452		
		Conc. (ng/mL)	0	0.049		94.857	13.339
	Matrix 2	PAR	0.009	0.099	11.099		
		Conc. (ng/mL)	0	0.059		114.87	4.757
	Matrix 3	PAR	0.018	0.097	5.529		
		Conc. (ng/mL)	0.001	0.055		107.666	6.630
	Matrix 4	PAR	0.029	0.096	3.338		
		Conc. (ng/mL)	0.003	0.053		104.159	14.469
	Matrix 5	PAR	0.007	0.092	14.098		
		Conc. (ng/mL)	0	0.049		95.3	16.254
	Matrix 6	PAR	0.017	0.09	5.449		
		Conc. (ng/mL)	0.001	0.046		89.313	5.889
	Mean	PAR	0.015	0.094	**8.161**		
		Conc. (ng/mL)	0.001	0.052		**101.028**	**10.223**
ERECVEETCSY							
	Matrix 1	PAR	0.077	0.151	1.954		
		Conc. (ng/mL)	0.186	1.122		87.693	3.196
	Matrix 2	PAR	0.041	0.158	3.87		
		Conc. (ng/mL)	0.035	1.324		103.4	15.848
	Matrix 3	PAR	0.029	0.155	5.343		
		Conc. (ng/mL)	0.013	1.196		93.439	3.548
	Matrix 4	PAR	0.027	0.149	5.507		
		Conc. (ng/mL)	0.01	1.144		89.411	19.678
	Matrix 5	PAR	0.017	0.173	10.341		
		Conc. (ng/mL)	0.003	1.628		127.194	1.988
	Matrix 6	PAR	0.051	0.152	2.99		
		Conc. (ng/mL)	0.071	1.156		90.329	9.828
	Average	PAR	0.04	0.147	**5.001**		
		Conc. (ng/mL)	0.053	0.051		**98.578**	9.014

PAR, peak area ratio.

[a] Calculated by dividing the peak area ratio of calculator 1 (LLOQ concentration) by the zero sample.

[b] Recovery of assay (%) = back-calculated concentration at the calibrator 1 (LLOQ concentration) × 100.

3.2.4. Carryover

The mean carryover of the analyte in the blank sample was 6.1% and 3.6% of the LLOQ concentrations for GYQELLEK and VDFTEIQK, respectively. The mean peak area of internal standard was 5.0% and 3.0% in the blank samples for GYQELLEK and VDFTEIQK, respectively. For ERECVEETCSY, the average carryover in the analyte channel and internal standard (IS) channel was 4.53% and 0.08%, respectively. No relationship was observed between the carryover and the analyte concentration (Table 14).

Table 14. Carryover measurements using blank sample run after the highest calibrator.

Peptide	Matrix	Peptide type[a]	Peak area of LLOQ (calibrator 1)	Peak area of blank sample[b]	Carryover (%)[c]
GYQELLEK		unlabeled			
	Matrix 1		52.000	1622.333	3.205
	Matrix 2		79.000	6156.667	1.283
	Matrix 3		77.000	1471.333	5.233
	Matrix 4		75.333	831.333	9.062
	Matrix 5		83.667	1248.667	6.700
	Matrix 6		83.000	1834.333	4.525
	Mean		**75.000**	**2194.111**	**5.001**
		labeled			
	Matrix 1		1.000	233.333	0.429
	Matrix 2		1.667	896.000	0.186
	Matrix 3		13.667	219.000	6.240
	Matrix 4		20.667	117.000	17.664
	Matrix 5		16.667	185.667	8.977
	Matrix 6		8.333	262.000	3.181

	Mean	**10.333**	**318.833**	**6.113**
VDFTEIQK	unlabeled			
	Matrix 1	122.333	1624.667	7.530
	Matrix 2	49.000	2785.333	1.759
	Matrix 3	65.333	5438.000	1.201
	Matrix 4	93.000	2049.000	4.539
	Matrix 5	101.000	6989.667	1.445
	Matrix 6	77.333	4972.333	1.555
	Mean	**84.667**	**3976.500**	**3.005**
	labeled			
	Matrix 1	5.333	155.333	3.433
	Matrix 2	13.667	275.000	4.970
	Matrix 3	2.000	528.000	0.379
	Matrix 4	13.667	199.333	6.856
	Matrix 5	9.000	640.667	1.405
	Matrix 6	21.000	449.000	4.677
	Mean	**10.778**	**374.556**	**3.620**
ERECVEETCSY	unlabeled			
	Matrix 1	1997.333	21.333	1.068
	Matrix 2	1383.333	25.333	1.831
	Matrix 3	953.333	57.667	6.049
	Matrix 4	640.000	17.667	2.760
	Matrix 5	1405.000	9.000	0.641
	Matrix 6	540.333	80.000	14.806
	Mean	**1153.222**	**35.167**	**4.526**
	labeled			
	Matrix 1	17605.333	6.333	0.036
	Matrix 2	12379.333	0.000	0.000
	Matrix 3	8442.667	14.667	0.174
	Matrix 4	5490.333	4.667	0.085
	Matrix 5	12476.333	3.667	0.029
	Matrix 6	4824.000	7.333	0.152
	Mean	**10203.000**	**6.111**	**0.079**

[a] Stable isotope-unlabeled peptide as internal standard and stable isotope-labeled peptide as target analyte.

[b] For blank sample, the peak area was from endogenous peptide or interferents corresponding to m/z of the designated peptides.

[c] Calculated by dividing the peak area of the blank sample by that of the LLOQ (calibrator 1) and multiplying the result by 100.

3.2.5. Precision and Recovery of assay

For GYQELLEK and VDFTEIQK, the CV_{intra} values were 3.4% to 7.0% and 2.8% to 10.3%, respectively, and the recovery of intra-assay were 90.3% to 116.0% and 90.3% to 101.9% for GYQELLEK and VDFTEIQK, respectively. The CV_{inter} values were 3.0% to 11.2% and 4.0% to 5.0% for GYQELLEK and VDFTEIQK, respectively, whereas the recovery of inter-assay were 88.4% to 107.2% and 86.9% to 101.4% for GYQELLEK and VDFTEIQK, respectively. Thus, the MRM-MS assay had satisfactory precision and recovery of assay within the established concentration range of AFP and AFP-L3 in human serum samples (Table 15).

For ERECVEETCSY, the CV_{intra} values were 6.1% to 10.4%, and the $accuracy_{intra}$ values were 92.8% to 112.7%, whereas the CV_{inter} values were 4.8% to 18.0% and the $accuracy_{inter}$ values were 84.2% to 120.7% for ERECVEETCSY. Thus, the SRM-MS assay had satisfactory accuracy and precision in the established concentration range of PIVKA-II (Table 15).

Table 15. Precision and recovery of assay measured in 4 QC samples (lower limit of quantification, low, medium, and high) over 6 days.

Peptide		Measurements	QC1	QC2	QC3	QC4
GYQELLEK		Expected conc. (ng/mL)	0.051	0.154	2000.256	3600
	Intra-assay[a]	Conc. (ng/mL)	0.059	0.169	1937.411	3249.274
		SD	0.003	0.012	136.23	109.457
		CV_{intra} (%)	**5.149**	**6.893**	**7.032**	**3.369**
		Recovery of assay (%)[c]	**115.979**	**109.899**	**96.869**	**90.258**
	Inter-assay[b]	Conc. (ng/mL)	0.055	0.159	2015.015	3184.032
		SD	0.006	0.009	187.764	94.294
		CV_{inter} (%)	**11.218**	**5.865**	**9.318**	**2.961**
		Recovery of assay (%)[c]	**107.21**	**103.507**	**100.749**	**88.445**
		CV_{total} (%)[d]	**12.343**	**9.051**	**11.674**	**4.485**
V*D*FTEIQK	Intra-assay[a]	Conc. (ng/mL)	0.046	0.141	1955.724	3668.717
		SD	0.005	0.006	71.414	101.406
		CV_{intra} (%)	**10.28**	**4.078**	**3.652**	**2.764**
		Recovery of assay (%)[c]	**90.337**	**91.801**	**97.785**	**101.909**
	Inter-assay[b]	Conc. (ng/mL)	0.048	0.134	1912.751	3649.914
		SD	0.002	0.007	76.163	166.858
		CV_{inter} (%)	**4.353**	**5.048**	**3.982**	**4.572**
		Recovery of assay (%)[c]	**94.317**	**86.935**	**95.636**	**101.387**
		CV_{total} (%)[d]	**11.163**	**6.489**	**5.403**	**5.342**
ERECVEETCSY		Expected conc. (ng/mL)	1.28	3.84	50000.64	90000
	Intra-assay[a]	Conc. (ng/mL)	1.442	3.73	48503.387	93944.823
		SD	0.15	0.28	4541.805	5724.886
		CV_{intra} (%)	**10.405**	**7.501**	**9.364**	**6.094**
		Recovery of assay (%)[c]	**112.667**	**102.945**	**92.778**	**95.801**
	Inter-assay[b]	Conc. (ng/mL)	1.545	3.234	46566.074	97537.706
		SD	0.173	0.438	8366.292	4660.973
		CV_{inter} (%)	**11.168**	**13.541**	**17.966**	**4.779**

Recovery of assay (%)[c]	**120.683**	**84.23**	**103.479**	**108.375**
CV_{total} (%)[d]	**15.264**	**15.48**	**20.26**	**7.744**

[a] Calculated by averaging the mean values of the six replicates on each day over 6 d.

[b] Calculated by averaging the first replicate of each day over 6 d.

[c] Recovery of assay (%) = measured concentration/expected concentration × 100.

[d] $\text{Total CV} = \sqrt{\text{CV}_{\text{intra}}{}^2 + \text{CV}_{\text{inter}}{}^2}$

3.2.6. Matrix effect

The mean matrix effect for the calibrators was 74.9% to 104.0% for GYQELLEK, 92.7% to 113.2% for VDFTEIQK and 97.3% to 128.3% for ERECVEETCSY. The CV values for all calibrators were within 12.9%. For the QC samples, the mean matrix effect was 99.4% to 100.5% and 95.4% to 107.9% for GYQELLEK VDFTEIQK, and 99.5% to 102.9% for ERECVEETCS with CV values less than 15% (Table 16). For GYQELLEK and VDFTEIQK, the mean recovery of spiking were 87.3% to 113.7% and 92.2% to 105.6%, respectively (Table 17).

The measured concentrations of the unmixed samples and the ratio at which they were mixed were taken into account when the expected concentration of the mixed sample was calculated. The bias between the expected and measured values was within the range of values that was deemed to be satisfactory (recovery of assay < ± 15% of target concentration), ranging from 2.2–1665.7 ng/mL for GYQELLEK and 12.9–712.1 ng/mL for VDFTEIQK. The mixed samples showed good linearity, as evidenced by their high coefficient of determination ($R2 > 0.99$) by first-order polynomial (straight line) regression (Figure 18 and Table 18).

The concentrations of GYQELLEK and VDFTEIQK were unaffected by icteric samples, but affected by lipemia. The concentrations of the two peptides differed in that VDFTEIQK was affected by hemolysis, as opposed to GYQELLEK which remained unaffected. Also, ERECVEETCSY was unaffected by lipemic or icteric samples but was influenced by hemolysis (Table 19).

The concentrations of AFP, AFP-L3 and PIVKA-II were similar between serum and plasma samples. The AFP concentration (ng/mL) in serum (plasma) was 933.26 (911.44), 252.59 (267.01), 13.14 (11.63), 127.86 (113.99), 540.32 (495.02), and 45.28 (44.21). The AFP-L3% values in serum (plasma) were 3.34 (3.52), 15.49 (14.40), 29.37 (33.08), 32.38 (34.05), 4.17 (3.83), and 17.44 (16.37) (Figure 19A). The PIVKA-II concentrations (ng/mL) in serum (plasma) were 60.5 (134.7), 753.5 (1018.1), 824.9 (1126.5), 2636.4 (2237.7), 109,340.7 (112497.1), and 49.1 (51.5). The concentrations of PIVKA-II were similar between serum and plasma samples (Figure 19B).

Table 16. Matrix effects of 6 matrices using 8 calibrators and 4 QC samples.

Peptide	Matrix	Measurements	Calibrator 1	Calibrator 2	Calibrator 3	Calibrator 4	Calibrator 5	Calibrator 6	Calibrator 7	Calibrator 8	QC 1	QC 2	QC 3	QC 4
GYQELLEK														
	Buffer matrix[a]	PAR	0.20	0.30	0.52	0.90	1.81	3.35	6.23	12.52	0.16	0.23	8.85	11.11
	Matrix 1[b]	PAR	0.15	0.35	0.46	0.93	1.92	3.37	6.31	13.20	0.15	0.22	9.21	11.04
		ME (%)[c]	75.41	118.96	88.99	103.44	106.25	100.56	101.39	105.42	97.76	96.48	104.07	99.32
	Matrix 2[b]	PAR	0.15	0.25	0.44	0.88	1.63	3.35	6.05	11.59	0.16	0.22	8.59	10.91
		ME (%)[c]	75.09	84.70	84.57	98.37	90.25	99.80	97.16	92.61	102.72	98.56	97.09	98.16
	Matrix 3[b]	PAR	0.15	0.32	0.50	0.95	1.96	3.38	6.00	12.10	0.16	0.22	8.50	10.90
		ME (%)[c]	75.92	107.60	97.29	105.97	108.59	100.72	96.34	96.63	100.05	97.28	96.02	98.11
	Matrix 4[b]	PAR	0.14	0.30	0.54	0.91	1.94	3.29	6.79	13.02	0.16	0.23	8.80	10.69
		ME (%)[c]	72.11	99.51	103.71	101.84	107.11	98.11	108.96	104.02	100.24	101.38	99.47	96.25
	Matrix 5[b]	PAR	0.15	0.30	0.56	0.95	1.84	3.25	6.42	12.47	0.15	0.23	9.04	10.66
		ME (%)[c]	78.11	102.72	107.44	105.63	102.05	96.83	103.13	99.60	98.80	100.68	102.17	95.90
	Matrix 6[b]	PAR	0.14	0.28	0.51	0.97	1.71	3.25	5.97	12.70	0.16	0.23	8.80	10.74
		ME (%)[c]	72.72	94.66	99.32	108.61	94.85	96.79	95.94	101.41	103.34	101.95	99.42	96.64
		Mean	**74.89**	**101.36**	**96.89**	**103.98**	**101.52**	**98.80**	**100.49**	**99.95**	**100.48**	**99.39**	**99.71**	**97.40**
	ME (%)[c]	SD	2.20	11.63	8.69	3.59	7.42	1.80	5.07	4.77	2.18	2.27	3.02	1.34
		CV (%)	**2.94**	**11.47**	**8.97**	**3.45**	**7.31**	**1.82**	**5.05**	**4.77**	**2.17**	**2.29**	**3.03**	**1.37**
VDFTEIQK														
	Buffer matrix[a]	PAR	0.10	0.17	0.36	0.58	0.93	2.14	4.10	7.32	0.10	0.15	5.23	6.51
	Matrix 1[b]	PAR	0.09	0.22	0.32	0.67	1.06	2.34	4.53	7.99	0.10	0.14	5.66	7.46
		ME (%)[c]	91.41	132.92	89.18	115.89	113.76	109.64	110.41	109.15	98.41	94.26	108.27	114.50
	Matrix 2[b]	PAR	0.10	0.19	0.34	0.52	1.08	2.17	4.14	7.72	0.10	0.14	5.43	6.96
		ME (%)[c]	98.55	113.74	94.32	90.69	116.31	101.77	100.96	105.43	98.01	96.44	103.85	106.89
	Matrix 3[b]	PAR	0.10	0.17	0.36	0.64	1.03	2.23	4.37	7.50	0.10	0.14	5.44	7.01

		ME (%)[c]	96.09	102.56	98.19	109.90	110.92	104.64	106.43	102.43	100.58	94.18	104.02	107.53
	Matrix 4[b]	PAR	0.10	0.19	0.32	0.61	1.06	2.39	4.32	7.82	0.10	0.14	5.41	7.08
		ME (%)[c]	94.74	112.88	89.28	105.35	114.48	111.98	105.31	106.82	98.79	96.24	103.45	108.64
	Matrix 5[b]	PAR	0.09	0.18	0.35	0.62	1.06	1.87	3.70	6.77	0.10	0.14	5.44	6.93
		ME (%)[c]	91.51	107.46	95.68	107.96	113.94	87.62	90.20	92.46	98.33	96.67	104.01	106.34
	Matrix 6[b]	PAR	0.09	0.15	0.32	0.56	1.02	1.88	3.68	6.60	0.10	0.14	5.32	6.75
		ME (%)[c]	89.40	90.14	89.78	97.07	109.72	87.92	89.70	90.24	98.87	94.34	101.69	103.58
		Mean	**93.62**	**109.95**	**92.74**	**104.47**	**113.19**	**100.59**	**100.50**	**101.09**	**98.83**	**95.36**	**104.22**	**107.91**
	ME (%)[c]	SD	3.43	14.17	3.85	9.14	2.43	10.57	8.71	7.88	0.91	1.21	2.18	3.64
		CV (%)	**3.66**	**12.89**	**4.16**	**8.75**	**2.15**	**10.50**	**8.67**	**7.80**	**0.92**	**1.27**	**2.09**	**3.38**
ERECVEETCSY														
	Buffer matrix[a]	PAR	0.12	0.28	0.49	0.92	1.62	2.71	5.17	9.18	0.16	0.23	6.69	8.73
	Matrix 1[b]	PAR	0.15	0.28	0.50	0.91	1.58	2.74	5.09	8.84	0.16	0.23	6.82	8.56
		ME (%)[c]	124.11	97.94	102.38	98.56	97.79	100.88	98.29	96.36	96.40	103.96	102.07	98.00
	Matrix 2[b]	PAR	0.16	0.27	0.51	0.91	1.58	2.70	5.05	8.93	0.17	0.24	6.82	8.71
		ME (%)[c]	129.77	95.58	104.29	99.22	97.70	99.59	97.64	97.37	103.64	107.37	102.04	99.79
	Matrix 3[b]	PAR	0.15	0.28	0.51	0.91	1.65	2.67	5.05	8.91	0.16	0.24	6.88	8.73
		ME (%)[c]	126.94	98.42	103.00	98.93	101.88	98.45	97.62	97.07	97.41	104.75	102.97	100.00
	Matrix 4[b]	PAR	0.15	0.28	0.52	0.91	1.58	2.91	5.00	8.63	0.16	0.22	6.77	8.67
		ME (%)[c]	122.20	99.37	105.99	99.24	97.54	107.01	96.57	94.05	100.48	99.34	101.30	99.26
	Matrix 5[b]	PAR	0.17	0.28	0.50	0.99	1.57	2.91	4.90	8.91	0.17	0.23	6.88	8.86
		ME (%)[c]	141.82	98.34	100.89	107.86	97.04	107.14	94.79	97.06	103.22	99.61	102.97	101.52
	Matrix 6[b]	PAR	0.15	0.28	0.52	0.93	1.61	2.89	5.11	8.77	0.16	0.23	6.84	8.57
		ME (%)[c]	124.78	99.84	106.68	100.74	99.63	106.60	98.82	95.54	98.55	102.28	102.23	98.18
		Average	**128.27**	**98.25**	**103.87**	**100.76**	**98.60**	**103.28**	**97.29**	**96.24**	**99.95**	**102.89**	**102.26**	**99.46**
	ME (%)[c]	**STDEV**	7.13	1.49	2.21	3.56	1.83	4.06	1.44	1.26	3.02	3.11	0.64	1.30
		CV (%)	**5.56**	**1.51**	**2.13**	**3.53**	**1.86**	**3.93**	**1.48**	**1.31**	**3.02**	**3.02**	**0.62**	**1.31**

PAR, peak area ratio; ME, matrix effect, SD, standard deviation; CV, coefficient of variation; QC, quality control.

[a] Stable isotope-unlabeled and labeled protein analogs spiked into elution buffer alone at the beginning of assay.

[b] Stable isotope-unlabeled and labeled protein analogs spiked into serum at the beginning of assay.

[c] Matrix effect = (serum matrixb / buffer matrixa) × 100 (%).

Table 17. Evaluation of recovery after spike.

Peptide	Matrix	Measurements	Zero sample	QC 1	QC 2	QC 3	QC 4
		Expected conc. (ng/mL)		0.051	0.154	2000.256	3600.000
GYQELLEK							
	Matrix 1	Conc. (ng/mL)	0.002	0.056	0.139	2078.250	3302.442
		Recovery of spiking (%)[a]		105.504	89.603	103.911	91.734
	Matrix 2	Conc. (ng/mL)	0.002	0.063	0.147	1738.242	3204.770
		Recovery of spiking (%)[a]		120.503	94.961	86.911	89.021
	Matrix 3	Conc. (ng/mL)	0.000	0.059	0.142	1698.987	3203.293
		Recovery of spiking (%)[a]		112.910	91.658	84.948	88.980
	Matrix 4	Conc. (ng/mL)	0.000	0.059	0.158	1858.883	3046.874
		Recovery of spiking (%)[a]		112.779	102.093	92.943	84.635
	Matrix 5	Conc. (ng/mL)	0.000	0.057	0.155	1984.921	3019.497
		Recovery of spiking (%)[a]		108.507	100.167	99.245	83.875
	Matrix 6	Conc. (ng/mL)	0.001	0.064	0.160	1851.968	3079.762
		Recovery of spiking (%)[a]		122.091	103.358	92.597	85.549
	Recovery of spiking (%)[a]	Mean		**113.716**	**96.973**	**93.426**	**87.299**
		SD		6.516	5.723	7.198	3.077
		CV (%)		5.730	5.902	7.704	3.525
V*D*FTEIQK							
	Matrix 1	Conc. (ng/mL)	0.000	0.054	0.144	2033.053	4149.100
		Recovery of spiking (%)[a]		104.255	93.518	101.651	115.253

Matrix 2	Conc. (ng/mL)	0.000	0.053	0.153	1824.947	3475.012
	Recovery of spiking (%)[a]		103.091	99.639	91.246	96.528
Matrix 3	Conc. (ng/mL)	0.001	0.057	0.144	1835.920	3533.395
	Recovery of spiking (%)[a]		110.598	93.655	91.795	98.150
Matrix 4	Conc. (ng/mL)	0.003	0.054	0.152	1807.934	3625.074
	Recovery of spiking (%)[a]		105.699	98.914	90.396	100.696
Matrix 5	Conc. (ng/mL)	0.000	0.054	0.154	1835.176	3436.746
	Recovery of spiking (%)[a]		104.165	100.171	91.758	95.465
Matrix 6	Conc. (ng/mL)	0.001	0.054	0.145	1729.486	3206.303
	Recovery of spiking (%)[a]		105.826	94.001	86.473	89.064
Spike recovery (%)[a]	Mean		**105.606**	**96.650**	**92.220**	**99.193**
	SD		2.653	3.233	5.033	8.773
	CV (%)		2.513	3.345	5.457	8.845

QC, quality control; SD, standard deviation; CV, coefficient of variation.

a Recovery of spiking (%) = $\frac{\text{(measured concentration of the spiked zero sample - measured concentration of the unspiked zero sample)}}{\text{expected concentration of the spiked zero sample}}$ X 100.

Table 18. Linearity evaluation using matrix mixed samples.

Peptide	Mixing ratio	Measured conc. (ng/mL)[a]	Expected conc. (ng/mL)[b]	Bias (%)[c]
GYQELLEK	A : B			
	100 : 0	2.231		
	80 : 20	330.258	334.921	**-1.392**
	60 : 40	695.181	667.611	**4.130**
	40 : 60	964.406	1000.300	**-3.588**
	20 : 80	1217.663	1332.990	**-8.652**
	0 : 100	1665.680		
V*D*FTEIQK	C : D			
	100 : 0	12.861		
	80 : 20	131.185	152.704	**-14.092**
	60 : 40	303.211	292.547	**3.645**
	40 : 60	404.823	432.389	**-6.375**
	20 : 80	522.771	572.232	**-8.644**
	0 : 100	712.075		

[a] Calculated by mean value of 6 replicates.

[b] Calculated by measuring the concentration of 2 unmixed samples and applying the ratio of the mix of the 2 samples.

[c] Bias (%) = $\frac{\text{(measured concentration - expected concentration)}}{\text{expected concentration}}$

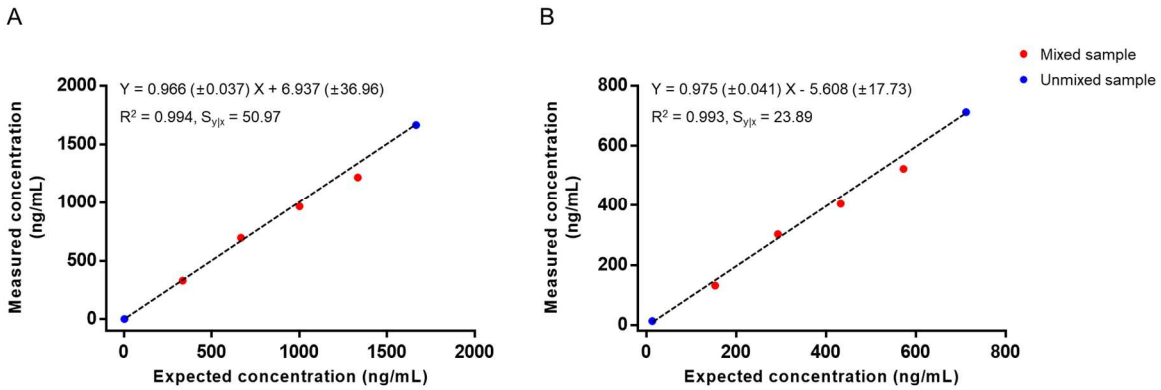

Figure 18. Matrix mixing experiment results.

To assess linearity, 2 HCC serum samples were mixed at various ratios (non-serial dilution) and analyzed by MRM-MS. The linearity of GYQELLEK (A) and V*D*FTEIQK (B) was evaluated by mixing serum A (high AFP) with B (low AFP) and C (high AFP-L3) with D (low AFP-L3) in 6 proportions (100:0, 80:20, 60:40, 40:60, 20:80, and 0:100), based on the concentrations as measured by LiBA. First-order polynomial regression was applied to evaluate linearity. Measurements of GYQELLEK (representing AFP) and V*D*FTEIQK (representing AFP-L3) in the 6 samples demonstrated a strong linear fit, as evidenced by R^2 (coefficient of determination) > 0.99. The dotted line represents the line of identity (y=x). $S_{y|x}$, standard deviation of residual.

Table 19. Matrix effects of endogenous interferents.

Normal serum and problem samples	GYQELLEK		Bias (%)[a]	V*D*FTEIQK		Bias (%)[a]	ERECVEETCSY		Bias (%)[a]
	Measured conc. (ng/mL)	Expected conc. (ng/mL)		Measured conc. (ng/mL)	Expected conc. (ng/mL)		Measured conc. (ng/mL)	Expected conc. (ng/mL)	
"Normal" serum	192.945			103.78			303.3		
Lipemic sample	12.308			1.381			2.35		
Mix 1 to 1 ("Normal" serum and Lipemic sample)	80.284	102.626	**-21.77**	43.019	52.58	**-18.183**	149.82	152.82	**1.97**
Hemolyzed sample	67.542			8.845			16.67		
Mix 1 to 1 ("Normal" serum and Hemolyzed sample)	115.199	130.244	**-11.551**	44.651	56.312	**-20.707**	203.04	159.99	**-26.91**
Icteric sample	27.125			3.925			6.38		
Mix 1 to 1 ("Normal" serum and Icteric sample)	114.749	110.035	**4.284**	53.573	53.852	**-0.518**	144.78	154.84	**6.5**

[a] Bias (%) $= \frac{\text{(measured concentration - expected concentration)}}{\text{expected concentration}} \text{x100}$.

Underline indicates that the value did not meet the criteria.

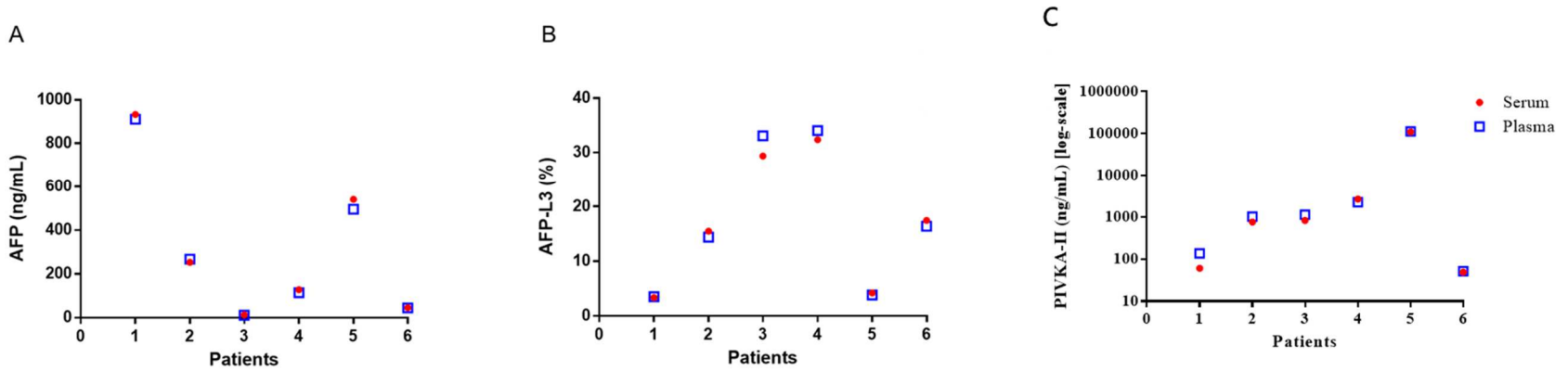

Figure 19. Evaluation of matrix effects from exogenous interferents.

AFP and AFP-L3 concentrations in samples collected in serum separation tubes (SST) and tri-potassium-ethylenediaminetetraacetic acid (K3-EDTA) tubes. Pairs (serum-plasma) of samples were obtained from the same HCC patients. The concentrations of AFP (ng/mL) (A), AFP-L3 (%) (B) and PIVKA-II (C) in serum showed high similarity compared with those in plasma.

3.2.7. Recovery of immunoprecipitation

The total recovery of immunoprecipitation were 99.3% to 102.8% for GYQELLEK and 95.9% to 103.9% for VDFTEIQK (Table 20). The total immunoprecipitation recovery rates were 99.3% to 100.6% for ERECVEETCSY (Table 20) indicating that the immunoprecipitation steps did not alter the performance of the assay.

Table 20. Recovery after immunoprecipitation.

Peptide	Matrix	Measurements	QC 1	QC 2	QC 3	QC 4
GYQELLEK						
	Matrix 1	PAR[a]	0.152	0.217	9.208	11.035
		PAR[b]	0.152	0.225	8.713	10.703
		Recovery of IP (%)[c]	100.032	96.765	105.678	103.101
	Matrix 2	PAR[a]	0.16	0.222	8.59	10.906
		PAR[b]	0.145	0.225	8.324	10.844
		Recovery of IP (%)[c]	109.786	98.762	103.195	100.568
	Matrix 3	PAR[a]	0.156	0.219	8.496	10.9
		PAR[b]	0.152	0.231	8.756	10.768
		Recovery of IP (%)[c]	102.666	94.897	97.025	101.224
	Matrix 4	PAR[a]	0.156	0.229	8.801	10.694
		PAR[b]	0.155	0.219	8.736	10.172
		Recovery of IP (%)[c]	100.633	104.174	100.751	105.126
	Matrix 5	PAR[a]	0.154	0.227	9.04	10.655
		PAR[b]	0.155	0.224	8.724	10.381
		Recovery of IP (%)[c]	98.949	101.238	103.624	102.633
	Matrix 6	PAR[a]	0.161	0.23	8.797	10.737
		PAR[b]	0.153	0.23	8.669	10.568
		Recovery of IP (%)[c]	104.794	99.937	101.466	101.599
	IP recovery (%)[c]	Mean	**102.81**	**99.295**	**101.956**	**102.375**
		SD	4	3.287	2.973	1.634
		CV (%)	3.89	3.31	2.915	1.596
V*D*FTEIQK						
	Matrix 1	PAR[a]	0.096	0.14	5.66	7.459

		PAR[b]	0.094	0.14	5.388	6.66
		Recovery of IP (%)[c]	102.453	100.113	105.063	112.004
	Matrix 2	PAR[a]	0.095	0.144	5.429	6.963
		PAR[b]	0.099	0.148	5.331	6.486
		Recovery of IP (%)[c]	96.613	96.946	101.845	107.36
	Matrix 3	PAR[a]	0.098	0.14	5.438	7.005
		PAR[b]	0.097	0.148	5.352	6.694
		Recovery of IP (%)[c]	100.846	95.011	101.614	104.64
	Matrix 4	PAR[a]	0.096	0.143	5.408	7.077
		PAR[b]	0.098	0.153	5.369	7.042
		Recovery of IP (%)[c]	98.081	93.748	100.721	100.506
	Matrix 5	PAR[a]	0.096	0.144	5.438	6.927
		PAR[b]	0.099	0.149	5.32	6.899
		Recovery of IP (%)[c]	96.796	96.79	102.217	100.416
	Matrix 6	PAR[a]	0.096	0.141	5.316	6.747
		PAR[b]	0.097	0.152	5.492	6.845
		Recovery of IP (%)[c]	99.281	92.699	96.803	98.577
	IP recovery (%)[c]	Mean	**99.012**	**95.885**	**101.377**	**103.917**
		SD	2.315	2.657	2.68	5.103
		CV (%)	2.339	2.771	2.644	4.91
ERECVEETCSY						
	Matrix 1	PAR[a]	0.16	0.23	6.82	8.56
		PAR[b]	0.16	0.23	6.88	8.6
		Recovery of IP (%)[c]	**97.37**	**101.68**	**99.21**	**99.5**
	Matrix 2	PAR[a]	0.17	0.24	6.82	8.71
		PAR[b]	0.16	0.25	6.95	8.72
		Recovery of IP (%)[c]	**106.24**	**98.12**	**98.2**	**99.92**
	Matrix 3	PAR[a]	0.16	0.24	6.88	8.73
		PAR[b]	0.17	0.24	6.7	8.73
		Recovery of IP (%)[c]	**93.7**	**100.69**	**102.72**	**99.98**
	Matrix 4	PAR[a]	0.16	0.22	6.77	8.67
		PAR[b]	0.16	0.24	6.61	8.62
		Recovery of IP (%)[c]	**101.85**	**95.04**	**102.48**	**100.5**
	Matrix 5	PAR[a]	0.17	0.23	6.88	8.86
		PAR[b]	0.17	0.22	6.8	8.48
		Recovery of IP (%)[c]	**100.68**	**100.11**	**101.31**	**104.59**
	Matrix 6	PAR[a]	0.16	0.23	6.84	8.57
		PAR[b]	0.16	0.23	6.86	8.72
		Recovery of IP (%)[c]	**102.26**	**99.86**	**99.68**	**98.32**
		Mean	**100.35**	**99.25**	**100.6**	**100.47**

	IP recovery (%)[c]	SD	4.33	2.37	1.85	2.15
		CV (%)	4.32	2.39	1.84	2.14

PAR, peak area ratio; IP, immunoprecipitation; SD, standard deviation; CV, coefficient of variation; QC, quality control;.

[a] Stable isotope-unlabeled and labeled protein analogs spiked into serum at the beginning of assay.

[b] Stable isotope-unlabeled and labeled protein analogs spiked during the elution step, immediately after immunoaffinity.

[c] Recovery of IP = (serum matrix[a] / serum matrix[b]) × 100.

3.2.8. Dilution integrity

The dilution-corrected concentrations were compared with the initial concentration (8000 ng/mL). Dilution factors up to 2500-fold had CV values that were < 10% and biases that were < 15% (Table 21). However, the bias of the samples that were diluted to 20000-fold exceeded 15%, due to excessive dilution. Thus, it is possible to accurately analyze clinical samples, the concentrations of which exceed the analytical range of measurement, by diluting up to 2500 times.

For PIVKA-II, the dilution-corrected concentrations were compared with the initial concentration (300000 ng/mL). Dilution factors up to 10000-fold had biases that were < 15%. Thus, it is possible to accurately analyze clinical samples wherein the analytical range of measurements exceeds the dynamic range of 10000-fold (Table 21).

Table 21. Evaluation of dilution integrity.

Peptide	Dilution factor	Measured concentration (ng/mL)	Dilution-corrected concentration (ng/mL)	Bias (%)[b]
GYQELLEK	Neat[a]		8000	
	5	1569.192	7845.96	**-1.925**
	150	52.974	7946.155	**-0.673**
	500	17.167	8583.497	**7.294**
	2500	3.662	9156.071	**14.451**
	20000	0.467	9341.369	**16.767**
V*D*FTEIQK	Neat[a]		8000	
	5	1521.435	7607.177	**-4.91**
	150	54.182	8127.328	**1.592**
	500	17.305	8652.389	**8.155**
	2500	2.936	7340.158	**-8.248**
	20000	0.283	5666.076	**-29.174**
ERECVEETCSY				
	Neat[a]		300000	
	200	1496.52	299305	**-0.23**
	400	804.09	321638	**7.21**
	2000	167.63	335264	**11.75**
	10000	28.52	285216	**-4.93**

SD, standard deviation; CV, coefficient of variation.

[a] The value is above the upper limit of quantification

[b] Relative bias (%) = $\frac{(\text{dilution-corrected concentration} - \text{neat concentration})}{\text{neat concentration}} \times 100$.

Underline indicates that the value did not meet the criteria.

3.2.9. Stability

For AFP, AFP-L3, under long-term storage (-20 ℃ and -70 ℃) and after freeze-thaw cycles, the deviations from baseline values were within the acceptable limit (< ±15%) for the GYQELLEK and VDFTEIQK peptides. However, with regard to short-term stability (room temperature and 4 ℃), the recovery of assay for QC 1 and QC 2 did not meet the criteria (< 85%) (Figure 20 A~J). Considering these results, I concluded that samples should be analyzed within 1 month after collection if stored at -20 ℃ or -70 ℃ and within 4 d if stored at room temperature or 4 ℃ and that up to 7 freeze-thaw cycles is possible.

For PIVKA-II, under short-term storage at 4℃ and after freeze-thaw cycles, the deviations from baseline values were within the acceptable limits (< ±15%), and during short-term storage at room temperature, the samples were relatively stable, such that only the deviation of QC1 at Day 7 was less than 85%. After long-term storage at -20℃, it's the overall deviation from baseline increased, whereas for storage at -70℃, the overall deviation fell. In these cases, the deviations of QC1 exceeded 115% (Figure 21 A-E).

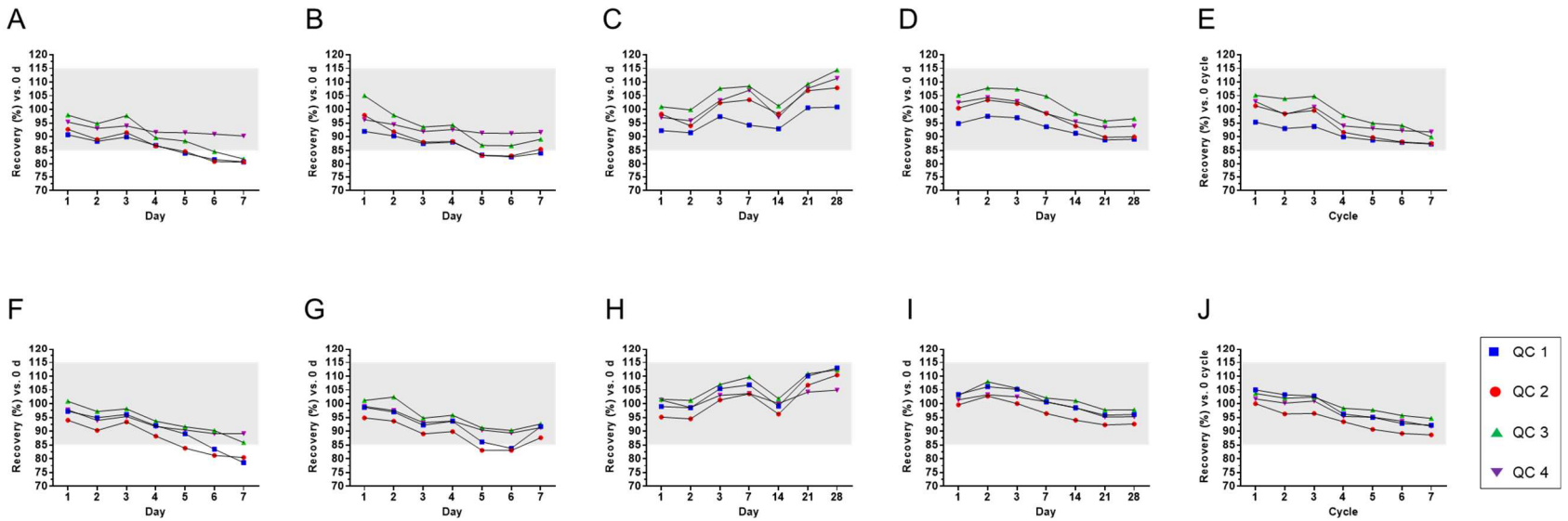

Figure 20. Evaluation of stability (GYQELLEK, VDFTEIQK).

The variation in GYQELLEK (A–E) and V*D*FTEIQK (F–J) concentrations compared with baseline concentrations (0 d or 0 cycle) during short-term storage at room temperature (A, F) and 4 °C (B, G); long-term storage at -20 °C (C, H) and 70 °C (D, I); and freeze-thaw cycles (E, J). The gray zone indicates the acceptable range (85% to 115%).

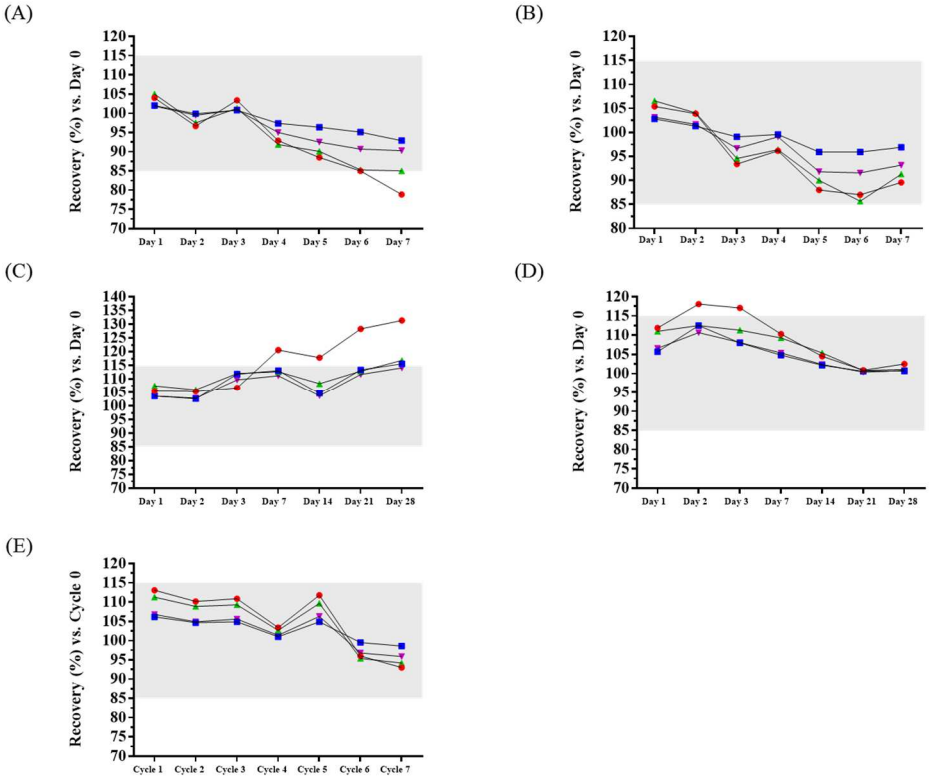

Figure 21. Evaluation of stability (ERECVEETCSY).

The variation in ERECVEETCSY concentrations compared with baseline concentration (Day 0 or Cycle 0) during short-term storage at (A) room temperature and (B) 4°C and long-term storage at (C) -20°C and (D) -70°C and (E) after a freeze-thaw cycle. The gray zone indicates the acceptable range (85% to 115%).

3.2.10. Reproducibility

The mean CV values of AFP and AFP-L3 concentrations over 6 d were within 15% for all samples, with the exception of AFP-L3 concentration of HCC1, which had a CV of 22.2%. These results demonstrate that the overall MRM-MS assay workflow was stable when using patient samples over several days (Figure 22A, B and Table 22)

The average CV values of the PIVKA-II concentration over 6 days were within 15% for all samples except for HCC1 sample, which had a CV of 17.6%. These results demonstrate that the overall SRM-MS assay workflow is suitable for patient samples over several days (Figure 22C and Table 22).

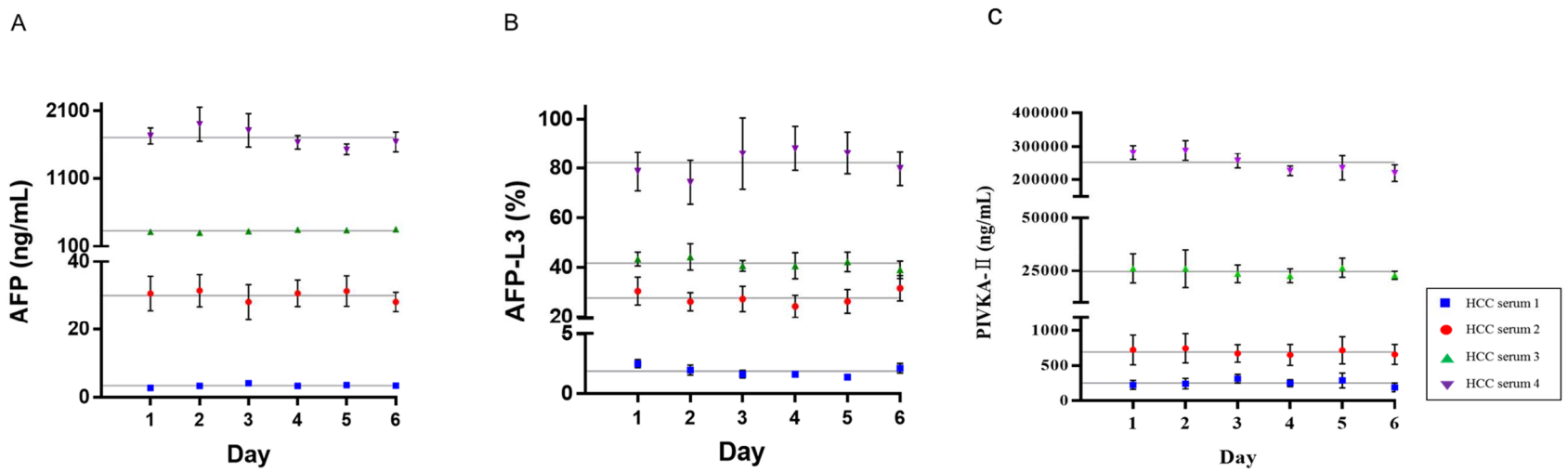

Figure 22. Evaluation of reproducibility.

Reproducibility of AFP (A), AFP-L3 (B) and PIVKA-II (C) was evaluated using 4 HCC serum patients. Injections were performed 6 times per day. Error bar represents the standard deviation of 6 replicates on a given day. Solid gray line marks the mean value over 6 d.

Table 22. Reproducibility of sample preparation over 6 days.

Analyte	Day	Measurements	HCC 1	HCC 2	HCC 3	HCC 4
AFP						
	1	Conc. (ng/mL)	2.781	30.659	315.163	1734.518
	2	Conc. (ng/mL)	3.378	31.495	299.949	1905.182
	3	Conc. (ng/mL)	4.132	28.172	321.880	1813.893
	4	Conc. (ng/mL)	3.333	30.715	343.657	1639.614
	5	Conc. (ng/mL)	3.596	31.363	335.790	1530.560
	6	Conc. (ng/mL)	3.448	28.168	352.042	1643.956
	Mean of 6 d	Conc. (ng/mL)	3.444	30.095	328.080	1711.287
		SD	0.437	1.529	19.354	134.904
		CV (%)	**12.675**	**5.079**	**5.899**	**7.883**
AFP-L3						
	1	Conc. (%)	2.500	30.366	43.288	78.678
	2	Conc. (%)	1.967	26.094	44.149	74.228
	3	Conc. (%)	1.620	27.165	40.571	85.990
	4	Conc. (%)	1.606	24.249	40.586	88.079
	5	Conc. (%)	1.369	26.241	42.150	86.221
	6	Conc. (%)	2.125	31.534	38.942	79.820
	Mean of 6 d	Conc. (%)	1.865	27.608	41.614	82.169
		SD	0.413	2.781	1.939	5.417
		CV (%)	**22.165**	**10.072**	**4.660**	**6.592**
PIVKA-II						
	1	Conc. (%)	226.480	725.200	26152.740	281016.160
	2	Conc. (%)	244.120	747.530	25997.400	287335.010
	3	Conc. (%)	315.130	675.080	23580.590	256255.920
	4	Conc. (%)	251.900	652.800	22611.540	225951.840
	5	Conc. (%)	289.710	717.900	26427.390	234582.840
	6	Conc. (%)	190.640	659.320	22854.610	218999.680
	Mean of 6 d	Conc. (%)	253.000	696.300	24604.040	250690.240
		SD	44.450	39.080	1774.410	28876.310
		CV (%)	**17.570**	**5.610**	**7.210**	**11.520**

HCC, hepatocellular carcinoma; AFP, alpha-fetoprotein; AFP-L3, lens culinaris agglutinin-reactive fraction of alpha-fetoprotein; SD, standard deviation; CV, coefficient of variation.

3.2.11. Quality control (QC) of samples and frequency

For AFP, AFP-L3, over 86.3% of all measurements were within ±15% for both peptides, and over 80.0% of QC samples at each concentration were within ±15%. Thus, both requirements were fulfilled, and the individual sample measurements were confirmed to be stable (Table 23).

For PIVKA-II, as shown in Table 2-18 over 95.0% of all measurements were biased within ±15%, and over 80% of QC samples at each level were biased within ±15%. All samples above the ±15% bias were at the LLOQ (QC1 samples).

Table 24 summarizes the results of the 12 items in the integrated multinational guidelines for validating analytical methods.

Table 23. Quality control results for individual samples.

GYQELLEK	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)
QC1				QC2			QC3			QC4		
	1	0.057	111.201	1	0.145	94.099	1	2149.185	107.458	1	3173.785	88.161
	2	0.052	101.285	2	0.135	87.744	2	2207.552	110.376	2	3546.109	98.503
	3	0.054	105.865	3	0.125	81.672	3	1895.846	94.791	3	3480.618	96.684
	4	0.062	121.295	4	0.143	92.793	4	2034.205	101.709	4	3142.9	87.303
	5	0.051	99.375	5	0.152	99.084	5	1831.074	91.553	5	3423.333	95.093
	6	0.066	128.056	6	0.156	101.759	6	1741.439	87.071	6	3367.137	93.532
	7	0.056	108.625	7	0.135	87.883	7	1910.16	95.507	7	3566.627	99.073
	8	0.056	108.82	8	0.119	77.415	8	1521.877	76.093	8	3119.247	86.646
	9	0.058	113.744	9	0.165	107.621	9	1808.689	90.433	9	3194.54	88.737
	10	0.059	114.932	10	0.158	102.905	10	2005.011	100.249	10	3237.326	89.926
	11	0.051	98.937	11	0.146	94.864	11	1661.462	83.072	11	3735.394	103.761
	12	0.056	108.814	12	0.149	96.909	12	1724.152	86.207	12	3317.664	92.157
	13	0.062	120.28	13	0.176	114.57	13	1733.371	86.667	13	3249.512	90.264
	14	0.057	110.827	14	0.152	99.153	14	1981.133	99.055	14	2948.288	81.897
	15	0.054	105.849	15	0.137	89.194	15	1983.721	99.185	15	2743.197	76.2
	16	0.054	104.515	16	0.153	99.419	16	2219.247	110.961	16	3086.348	85.732
	17	0.053	102.638	17	0.151	98.388	17	2200.026	110	17	3005.938	83.498
	18	0.058	113.351	18	0.155	100.594	18	1824.764	91.237	18	3158.735	87.743
	19	0.059	114.551	19	0.159	103.508	19	2161.221	108.06	19	3230.165	89.727
	20	0.066	128.963	20	0.146	95.358	20	1858.044	92.901	20	3364.928	93.47

V*D*FTEIQK	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)
QC1				**QC2**			**QC3**			**QC4**		
	1	0.057	111.928	1	0.135	88.038	1	2033.574	101.677	1	4089.295	113.592
	2	0.051	98.698	2	0.132	85.723	2	1985.129	99.255	2	4113.034	114.251
	3	0.055	106.621	3	0.161	104.874	3	2130.68	106.533	3	4082.592	113.405
	4	0.05	97.836	4	0.147	95.806	4	2035.184	101.758	4	4134.449	114.846
	5	0.05	98.487	5	0.138	89.928	5	2073.012	103.649	5	4220.214	117.228
	6	0.054	106.007	6	0.11	71.819	6	1837.709	91.884	6	3749.971	104.166
	7	0.051	98.643	7	0.165	107.511	7	1869.875	93.493	7	3596.04	99.89
	8	0.053	104.152	8	0.159	103.427	8	1767.861	88.392	8	3294.585	91.516
	9	0.056	109.753	9	0.162	105.188	9	1763.723	88.185	9	3430.255	95.285
	10	0.045	87.805	10	0.163	105.976	10	2080.976	104.047	10	3191.403	88.65
	11	0.065	127.36	11	0.163	105.876	11	1766.267	88.312	11	3832.257	106.452
	12	0.046	88.9	12	0.164	106.839	12	1902.029	95.1	12	3571.479	99.208
	13	0.062	122.016	13	0.138	90.137	13	1873.344	93.666	13	3742.986	103.972
	14	0.059	114.853	14	0.128	83.342	14	1844.286	92.213	14	3978.378	110.511
	15	0.048	93.157	15	0.163	105.977	15	1860.258	93.012	15	3829.066	106.363
	16	0.048	93.56	16	0.164	106.659	16	1900.503	95.024	16	3492.742	97.021
	17	0.053	102.668	17	0.141	91.775	17	1629.127	81.455	17	3179.917	88.331
	18	0.055	107.215	18	0.121	78.5	18	2006.467	100.322	18	3103.685	86.213
	19	0.051	100.029	19	0.134	86.957	19	1661.751	83.087	19	3133.763	87.049
	20	0.047	90.912	20	0.125	81.536	20	1705.429	85.27	20	3158.282	87.73

ERECVEETCSY	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)	Batch No.	Conc. (ng/mL)	Recovery of assay (%)
QC1				QC2			QC3			QC4		
	1	1.35	105.15	1	3.27	85.25	1	47,943.50	95.89	1	90713.56	100.79
	2	1.32	103	2	3.54	92.1	2	46,450.01	92.9	2	89416.25	99.35
	3	1.21	94.17	3	3.98	103.76	3	47,968.58	95.94	3	93300.75	103.67
	4	1.2	93.5	4	3.83	99.79	4	53,086.89	106.17	4	89081.77	98.98
	5	1.4	109.33	5	3.91	101.93	5	54,028.54	108.06	5	91725.43	101.92
	6	1.26	98.34	6	3.79	98.66	6	50,946.57	101.89	6	94260.47	104.73
	7	1.37	107.36	7	3.87	100.88	7	47,600.22	95.2	7	95236.02	105.82
	8	1.51	117.72	8	3.36	87.55	8	46,995.22	93.99	8	85950.19	95.5
	9	1.19	93.12	9	4.28	111.49	9	53,415.21	106.83	9	94044.02	104.49
	10	1.11	87.08	10	3.56	92.72	10	51,555.23	103.11	10	89815.95	99.8
	11	1.12	87.31	11	3.6	93.65	11	53,560.55	107.12	11	86696.55	96.33
	12	1.05	82.19	12	3.77	98.23	12	50,785.89	101.57	12	94129.26	104.59
	13	1.13	88.39	13	3.63	94.54	13	47,945.95	95.89	13	90267.53	100.3
	14	1.35	105.11	14	4.12	107.35	14	50,914.40	101.83	14	84763.73	94.18
	15	1.53	119.19	15	3.78	98.42	15	46,385.55	92.77	15	85488.53	94.99
	16	1.44	112.36	16	3.55	92.51	16	53,176.27	106.35	16	86006.58	95.56
	17	1.41	110.15	17	3.78	98.48	17	45,563.43	91.13	17	90621.51	100.69
	18	1.19	92.96	18	3.54	92.13	18	51,616.41	103.23	18	93982.8	104.43
	19	1.51	118.16	19	3.87	100.7	19	52,879.36	105.76	19	89598.88	99.55
	20	1.18	92.17	20	3.85	100.21	20	47,084.02	94.17	20	94818.5	105.35

A batch contained 20 patients and a QC set (QC1, QC2, QC3, and QC4); QC sets were analyzed after the patient samples.

Underline indicates that the value did not fulfill the criteria set by the guideline.

Table 24. Performance characteristics and results of the 12 items in the integrated multinational guidelines for validating analytical methods.

No.	Factor	Integrated guidelines		In this study	
		Validation practices	**Performance specification**	**Validation practices**	**Performance specification**
1	**Calibration curve**	**[FDA]** • Blank sample (matrix sample without analyte and internal standard), zero sample (matrix sample with internal standard only), and ≥ 6 calibrators should be prepared in the same biological matrix as the sample covering the expected range (including the LLOQ). **[EMA, KFDA]** • Same as FDA except that the expected range must include the ULOQ, as well. **[CLSI]** • Blank sample, zero sample, and 6–8 calibrators should be prepared in the same matrix as the samples. • Calibrators should cover the LLOQ, ULOQ, and any medically relevant decision point within the analytical measurement interval.	**[FDA, EMA, KFDA]** • Bias within 15% for all calibrators for (within 20% at the LLOQ) ≥ 75% of the calibrators. **[EMA]** • The back-calculated concentrations should be presented with the calculated mean accuracy. • At least 3 available (or acceptable) curves should be reported. **[EMA, KFDA]** • If replicates are used, the bias at each concentration must be < 15% in at least 50% of the replicates, except for the LLOQ (< 20%)**.** **[CLSI]** • Bias < 15% for all calibrators above the LLOQ. • $R^2 > 0.995$.	• Blank sample, zero sample, and 8 calibrators were prepared in 6 different matrices; each calibration curve was generated in the same matrix sample covering the LLOQ and ULOQ. • The 6 matrices were prepared daily and analyzed in 3 replicates over 6 d. • The LLOQ value is the lowest concentration, meeting precision (CV < 20%), recovery of assay* (within ± 15% bias), and S/N > 5. • The ULOQ value is the highest concentration, meeting precision (CV < 20%), and recovery of assay* (within ± 15% bias).	**(GYQELLEK, V*D*FTEIQK)** • 7 of 8 calibrators (87.5%) had a bias within 15%. **(GYQELLEK)** • Linearity (R^2) of reverse and forward curves was 0.998 and 0.988, respectively, from 0.051 (LLOQ) to 4000 ng/mL (ULOQ). **(V*D*FTEIQK)** • Linearity (R^2) of reverse and forward curves was 0.998 and 0.992, respectively, from 0.051 (LLOQ) to 4000 ng/mL (ULOQ).

2	**Analytical specificity**	**[FDA, EMA, KFDA]** • Each sample must be tested in at least 6 different blank matrices for interference. **[CLSI]** A matrix (preferably a patient sample) containing high concentrations of potential interferent should be evaluated with and without analyte.	**[FDA]** • Analytical specificity should be confirmed at the LLOQ. **[EMA, KFDA]** • Response from a potential interferent should be < 20% of the LLOQ for the analyte and < 5% for the IS. **[CLSI]** • If the signal of the qualifier ion is > 50% that of the quantifier ion, the ion ratio in the patient samples should be within 20% of the mean ratio of the standards.	• Compared the blank samples analyzed in 6 different matrices at their respective LLOQ concentrations. • The signal ratios (qualifier ion/quantifier ion) of all patient samples were compared with the mean ratio of the internal standard. • The matrix effect was also evaluated (results are described in the column below).	**(GYQELLEK)** • Potential interference responses were 2.0% for analyte and 3.2% for IS, on average. • Of the total patient samples, 2.0% exceeded the mean ratio of the internal standard by 20%. **(V*D*FTEIQK)** • Potential interference responses were 3.0% for analyte and 2.8% for IS, on average. • Of the total patient samples, 1.5% exceeded the mean ratio of the internal standard by 20%.
3	**Analytical sensitivity**	**[FDA]** • The LLOQ should be established using ≥ 5 samples. **[EMA]** • Intra-assay, the LLOQ should be established using ≥ 5 samples. • Inter-assay, the LLOQ should be established using samples from 3 runs on ≥ 2 different days. **[KFDA]** • Not addressed. **[CLSI]** • LLOQ should be established using ≥ 40 replicates from 3–5 different samples close to the predetermined limit of detection (LOD), over ≥ 5 runs.	**[FDA, EMA, KFDA]** • Signal-to-noise (S/N) > 5:1 at the LLOQ (noise response from a blank sample)**.** **[FDA, EMA, CLSI]** • Precision should be < 20% CV. **[FDA, EMA]** • Accuracy within 20%. **[CLSI]** • S/N > 20:1 at the LLOQ. • Accuracy within 15%.	• To determine the LLOQ, 5 different concentrations close to the LOD were analyzed in 3 replicates. • The LLOQ was established from 6 different matrix samples. • The 6 matrices were analyzed daily in 3 replicates over 6 d.	**(GYQELLEK, V*D*FTEIQK)** • LLOQ is determined at a concentration close to LOD. **(GYQELLEK)** • S/N was 10.8, precision was 7.7% CV, and recovery of assay* was 100.1% at the LLOQ determined in 6 different matrices. **(V*D*FTEIQK)** • S/N was 8.2, precision was 9.4% CV, and recovery of assay* was 101.0% at the LLOQ determined in 6 different matrices.

4	**Carryover**	**[FDA]** • Not addressed. **[EMA, KFDA]** • Assessed by injecting blank samples after a high-concentration sample or calibration standard at the ULOQ. **[CLSI]** • Assessed by injecting extracted negative samples after samples with increasing concentrations of analyte.	**[FDA]** • No specific performance criteria specified. **[EMA, KFDA]** • Carryover should be < 20% of the LLOQ and < 5% for the IS. **[CLSI]** • Carryover limit should be the highest concentration that does not carry over to a negative sample ≥ 25% of the LLOQ.	• Blank samples injected after the calibrator at the ULOQ. • Due to financial constraints (protein analogs), concentrations above 4000 ng/mL (ULOQ) were not evaluated.	**(GYQELLEK)** • Carryover was 6.1% for analyte and 5.0% for IS, on average. **(V*D*FTEIQK)** • Carryover was 3.6% for analyte and 3.0% for IS on average.
5	**Precision**	**[FDA]** • Assessed using ≥ 5 samples per concentration at ≥ 3 concentrations in analytical measurement interval. • Precision is subdivided into intra-assay (intra-batch precision or intra-assay repeatability) and inter-assay (inter-batch precision or inter-assay repeatability). **[EMA, KFDA]** • For intra-assay, precision should be assessed using ≥ 5 samples per concentration at LLOQ, low, medium, and high QC samples in a single run. • For inter-assay, precision should be assessed at 4 concentrations (LLOQ, low, medium, and high QC) from ≥ 3 runs analyzed on ≥ 2 different days. **[CLSI]** • Imprecision should be assessed using 20 samples at 3 concentrations	**[FDA, EMA, KFDA, CLSI]** • CV < 15% at all concentrations except for the LLOQ (< 20%). **[CLSI]** • No specific performance criteria specified (imprecision goals should be set based on predetermined total allowable error, biological variation, clinical guidelines by expert groups, and local or regional regulatory requirements)**.**	• For intra-assay, precision was assessed using 6 samples at 4 QCs [LLOQ, 3×LLOQ, (LLOQ + ULOQ)/2, 0.9×ULOQ]. • For inter-assay, precision was assessed using 4 QCs and analyzed from 6 replicates over 6 d.	**(GYQELLEK)** • Intra-assay CV in 4 QCs, 3.4–7.0%. • Inter-assay CV in 4 QCs, 3.0–11.2%. • Total CV in 4 QCs, 4.5–12.3%. **(V*D*FTEIQK)** • Intra-assay CV in 4 QCs, 2.8–10.3%. • Inter-assay CV in 4 QCs, 4.0–5.0%. • Total CV in 4 QCs, 5.3–11.2%.

		[110%×LLOQ, (LLOQ+ULOQ)/2, and 90%×ULOQ]. • Samples within 25% of any medical decision points should be evaluated			
6	**Recovery of assay***	**[FDA]** • Assessed using ≥ 5 samples of known analyte concentration at ≥ 3 concentrations (i.e., QC). **[EMA, KFDA]** • For intra-assay, accuracy should be assessed using ≥ 5 samples per concentration at ≥ 4 concentrations [LLOQ, < 3×LLOQ (low QC), 30-50% of the calibration curve range (medium QC), at least at 75% of the upper calibration curve range (high QC) in the analytical measurement interval] using spike and recovery analysis. • For inter-assay, accuracy should be assessed using 4 QC samples (LLOQ, low, medium, and high) from ≥ 3 runs analyzed on ≥ 2 different days. **[CLSI]** • Assessment of accuracy should include > 1 approach: (a) comparison to reference measurement procedures using ≥ 40 samples (preferably patient samples) covering the analytical measurement interval; (b) analysis of commutable certified reference materials; and/or (c) spike and recovery analysis.	**[FDA, EMA, KFDA, CLSI]** • The mean value < ±15% of the nominal value except for the LLOQ (< ±20%). **[CLSI]** • No specific requirement specified (acceptable criteria should be defined based on biological variation, clinical guidelines established by expert groups, and local or regional regulatory requirements).	• Recovery of Intra-assay* was assessed using 6 samples at 4 QCs [LLOQ, 3×LLOQ, (LLOQ + ULOQ)/2, 0.9×ULOQ]. • Recovery of Inter-assay* was assessed using 4 QCs analyzed daily in 6 replicates over 6 d. • The recovery of spiking and immunoprecipitation in 4 QCs is also evaluated (results are described in the column below).	**(GYQELLEK)** • Recovery of intra-assay* in 4 QCs, 90.3–116.0% (LLOQ < ± 20%). • Recovery of Inter-assay* in 4 QCs, 88.4–107.2%. **(V*D*FTEIQK)** • Recovery of intra-assay* in 4 QCs, 90.3–101.9%. • Recovery of Inter-assay* in 4 QCs, 86.9–101.4%.

7	**Matrix effect**	**[FDA]** • Ion suppression or enhancement should be addressed. **[EMA, KFDA]** • Assessed using ≥ 6 lots of blank matrix (native matrix) from individual samples (pooled matrix should not be used). • Matrix factor should be calculated for each lot of matrix by calculating the ratio of the peak area in the presence of matrix (measured by analyzing blank matrix spiked after extraction with analyte) to the peak area in absence of matrix (neat solution of the analyte). • Matrix effect should be determined at 2 specific concentrations (maximum of 3 times the LLOQ and close to the ULOQ). • If necessary, hemolyzed or lipemic samples should be tested**.** **[CLSI]** • Comparison of the peak area for ≥ 5 native matrix samples spiked with analyze postextraction analyte vs. analyte spiked into neat solution tested at all calibrators. **[CLSI - linearity]** • 9-11 concentrations should be analyzed with 2-4 replicates each. • Serial dilutions should be avoided.	**[FDA]** • No specific performance criteria specified. • Assessed to ensure that precision, selectivity, and analytical sensitivity will not be compromised. **[EMA, KFDA, CLSI]** • Precision CV < 15%. **[CLSI]** • Precision CV < 15%. • Evaluated in the context of total allowable error limits for specific analytes. **[CLSI - linearity]** • Any detected nonlinearity should be assessed for clinical significance.	• Peak area ratio for 6 different matrix samples spiked with analyte vs. analyte spiked into neat solution were compared at 8 calibrators and 4 QCs. • 6 concentrations of 2 samples mixed in different ratios (non-serial dilution) were analyzed in 6 replicates. Evaluated by first-order polynomial (straight line) regression method. • Evaluated the effects of lipemia, hemolysis, and icterus. • Evaluated the effects of 2 types (serum, plasma) of blood collection tubes.	**(GYQELLEK)** • CVs of all calibrators and QCs < 11.5%. • Recovery of spiking, 87.3–113.7%. • Linearity (R^2 = 0.994) was obtained at 2.2–1665.7 ng/mL. • Not affected by hemolysis and icterus but affected by lipemia. **(V*D*FTEIQK)** • CVs of all calibrators and QCs < 12.9%. • Recovery of spiking, 92.2–105.6%. • Linearity (R^2 = 0.993) was obtained at 12.9–712.1 ng/mL. • Not affected by icterus but affected by lipemia and hemolysis. **(GYQELLEK, V*D*FTEIQK)** • Similar between serum and plasma samples.

		• Evaluated with polynomials regression method.			
8	**Recovery of immunoprecipitation**	**[FDA, CLSI]** • Compare the analytical results for extracted samples at 3 concentrations (low, medium, and high) with unextracted standards, which represent 100% recovery. **[EMA]** • Not addressed. **[KFDA]** • Same as FDA but ≥ 3 replicates should be used at 3 concentrations (low, medium, and high). **[CLSI]** • When spiking endogenous analytes into matrices that are not blanks, standard addition should be used. • Proficiency testing, QC, and external reference samples should be included in recovery experiments.	**[KFDA]** • No specific performance criteria specified. **[EMA]** • No specific performance criteria specified. **[FDA, CLSI]** • Extent of recovery of an analyte and of the IS should be consistent, precise, and reproducible.	• The 4 QCs were prepared in 6 different matrix samples by comparing the results for enriched samples with those of unenriched samples, which represent 100% recovery.	**(GYQELLEK)** • Recovery of immunoprecipitation, 99.3–102.8%. **(VDFTEIQK)** • Recovery of immunoprecipitation, 95.9–103.9%.
9	**Dilution integrity**	**[FDA]** • Not addressed. **[EMA, KFDA]** • Assessed by spiking the matrix with an analyte concentration above the ULOQ and diluting this sample in a blank matrix with ≥ 5 samples per dilution factor. **[CLSI]** • ≥ 5 separate replicates of each intended dilution should be verified	**[FDA]** • No specific performance criteria specified. **[EMA, KFDA]** • Accuracy and precision within 15% for all dilutions. **[CLSI]** • At concentrations > 3 × the LLOQ, mean recoveries of 100% (SD < 15%) are acceptable; imprecision < 15% is acceptable.	• Samples were initially above the ULOQ (8000 ng/mL), and 5 dilution factors (5, 150, 500, 2500 and 20000) were used. • Each sample was analyzed in 6 replicates.	**(GYQELLEK)** • Up to 2500-fold dilution, precision and recovery of assay* were < 9.8% and < ± 14.5%, respectively. **(VDFTEIQK)** • Up to 2500-fold dilution, precision and recovery of assay* were < 7.7% and < ± 8.2%, respectively.

		(analyte-free native matrix is preferred for dilution).			
10	**Stability**	**[FDA, CLSI]** • At least 3 replicates at each of the low and high concentrations should be assessed. • Bench-Top stability should be conducted to cover laboratory handling conditions. • Long-term stability should equal or exceed the time between the date of first sample collection and the date of last sample analysis. • Stock solution stability should evaluate the stock solution of drug and internal standard. • Freeze and thaw stability: ≥ 3 cycles. • Processed sample stability should be determined, including the resident time in the autosampler. **[EMA, KFDA]** • At least 3 replicates at each of the low (3 × LLOQ) and high (close to the ULOQ) concentrations should be assessed. • Stability of the stock solution and working solutions. • Freeze and thaw stability. • Short-term stability (at room temperature, sample processing temperature). • Long-term stability (in the freezer).	**[FDA, EMA, KFDA]** • The mean concentration at each value should be within 15% of the nominal concentration. **[CLSI]** • No specific performance criteria specified (stability characteristics for samples, sample extracts, calibrators, and QC materials should be verified).	• Short-term stability (4 °C, and RT for 7 d), long-term stability (-20 °C, and -70 °C for 28 d), 7 freeze-thaw cycle experiments were done with 4 QC samples. • All samples were analyzed in 6 replicates.	**(GYQELLEK, V*D*FTEIQK)** • Long-term stability and freeze-thaw cycle test, 4 QCs are stable under all conditions. • Short-term stability test, QC1 and QC2 were unstable after 5 d.

		[EMA] • Regarding the stability of stock and working solutions, it is not needed to evaluate the stability of stable isotope-labeled internal standards if the isotope is demonstrated to be stable under the same conditions.			
11	**Reproducibility**	**[FDA]** • Reproducibility is assessed by replicate measurements, including QC samples and possibly incurred samples. **[EMA, KFDA, CLSI]** • Not addressed.	**[FDA, EMA, KFDA, CLSI]** • No specific performance criteria specified.	• 4 HCC samples were prepared daily for 6 d and analyzed in 6 replicates each day over 6 d.	**(AFP)** • CVs < 12.7% for all 4 HCC samples. **(AFP-L3)** • 3 of 4 HCC samples had CVs < 10.1% (22.2% for 1 sample).
12	**QC of samples and frequency**	**[FDA, KFDA]** • ≥ 3 QC concentrations (low, medium, and high) should be tested in duplicate. • Number of QC samples analyzed during a batch should represent ≥ 5% of the total number of unknown samples or six total QCs, whichever is greater. • QCs should be interspersed with study samples during processing and analysis. **[EMA]** • Not addressed. **[CLSI]** • ≥ 3 QC concentrations (3 × LLOQ, middle range, and near the ULOQ)	**[FDA, KFDA]** • At least 67% of the QCs ≤ ±15% of their nominal concentrations. • At least 50% of QCs at each concentration ≤ ±15% of their nominal concentrations. **[EMA]** • No specific performance criteria specified**.** **[CLSI]** • No specific performance criteria specified (laboratory must establish its own QC acceptability criteria).	• 4 QCs (LLOQ, low, medium, and high) were tested in singlicate. • Number of QC is 5% (4 QCs, 20 runs) of the total individual samples. • QC samples were interspersed with individual patient samples during analysis.	**(GYQELLEK)** • 86.3% of the QCs were ≤ ±15% of their nominal concentrations. **(V*D*FTEIQK)** • 87.5% of the QCs were ≤ ±15% of their nominal concentrations. **(GYQELLEK, V*D*FTEIQK)** • 80.0% of QCs at each concentration were ≤ ±15% of their nominal concentrations.

	tested in duplicate. • Number of QC samples analyzed during a batch should represent ≥ 5% of the total number of patient samples.

FDA, US Food and Drug Administration; EMA, European Medicines Agency; KFDA, Korea Food and Drug Administration; CLSI, Clinical and Laboratory Standards Institute; LLOQ, lower limit of quantification; ULOQ, upper limit of quantification; LOD, limit of detection; CV, coefficient of variance; S/N, signal-to-noise; IS, internal standard; QC, quality control; SD, standard deviation; AFP, alpha-fetoprotein; AFP-L3, lens culinaris agglutinin-reactive fraction of alpha-fetoprotein.

Content that does not meet the guidelines is underlined.

* Due to the lack of reference materials, I evaluated the accuracy using the measurement procedure “(c) spike and recovery analysis approach” in compliance with CLSI. Thus, the term "recovery of assay" was used instead of "accuracy".

IV. DISCUSSION

The US Food and Drug Administration (FDA) has cleared novel and specific HCC serum biomarkers, such as the lens culinaris agglutinin-reactive fraction of alpha-fetoprotein (AFP-L3), for assessing the risk of HCC *(6)*. AFP can be subfractionated into 3 distinct species–L1, L2, and L3–based on its reactivity to *lens culinaris agglutinin* (LCA) lectin, as evidenced by its migration pattern by affinity electrophoresis. The microheterogeneity of the glycan structure between AFP-L1 and AFP-L3 is attributed to the presence of an α-1,6 core fucose at the reducing end of the N-acetylglucosamine of AFP (7). AFP-L3 is a glycoform with core fucosylation. AFP-L1 is predominantly expressed in benign liver diseases, such as chronic hepatitis and liver cirrhosis. AFP-L2, which is produced by yolk sac tumors, has moderate affinity for LCA and is also detectable in maternal serum during pregnancy. AFP-L3 is tumor-specific for HCC. AFP-L3 concentrations correlate with AFP concentrations. However, AFP-L3% lacks such a relationship with AFP *(8, 9)*. As a biomarker, AFP-L3% is independent of AFP.

The first clinical laboratory assay for AFP-L3 was developed using a lectin affinity electrophoresis method. Subsequently, the ratio of L3 to L1, expressed as a percentage, was enabled to be determined using dye-labeled antibodies and quantified by densitometry. An automated assay was developed for clinical use in Japan on a liquid-phase binding assay (LiBA) platform. AFP-L3% values can be generated when AFP is > 10 ng/mL, with a minimal detectable limit of AFP-L3 of

0.8 ng/mL *(10)*. LiBA was cleared by the FDA for evaluating the risk of HCC in the US in April 2005. Since then, assay technologies have continued to evolve. Since 2009, such assays have been incorporated into microchip capillary electrophoresis and LiBA on a micro total analysis system (μTAS) Wako™i-30 autoanalyzer *(11)*. With the deployment of second-generation assays, the analytical sensitivity has improved by decreasing the minimal detection limit to 0.3 ng/mL of AFP-L3 *(12)*. The μTAS Wako™i-30 autoanalyzer obtained FDA clearance for in vitro diagnostic use in February 2011.

Although the clinical sensitivity of the assay has increased as its detection limit has improved, the general consensus of practicing clinicians is that the high incidence of false negatives renders the test results unreliable (24, 25). Accurate measurements of AFP-L3% have been limited to HCC patients with AFP concentrations greater than 0.3 ng/mL, due to the insufficient analytical sensitivity of the instrument. On the μTAS Wako™i-30 autoanalyzer, AFP-L3% cannot be reported if the total AFP level is < 0.3 ng/mL, even in cases of high AFP-L3 levels. The low sensitivity of AFP-L3 has hampered enthusiasm for its potential as an HCC-specific biomarker.

In this study, the AFP-L3-positive rates in 200 HCC samples were 61.5% (n = 123) and 81.0% (n = 162) by LiBA and MRM-MS assay, respectively–thus, the analytical sensitivity improved, and 39 additional HCC patients were observed. Further, the MRM-MS assay can identify more HCC patients with normal AFP levels (≤ 5.9 ng/mL; n = 82) than LiBA (n = 20, 24.4% versus n = 44, 53.7%). The MRM-

MS assay is not used for diagnostic decision-making, unlike LiBA. If the MRM-MS assay is applied in clinical practice, it will benefit certain HCC patients who have been designated only with the AFP concentration from LiBA measurements, wherein AFP-L3 measurements are available due to the greater sensitivity of MRM-MS assay. The cutoff value is 20 ng/mL for AFP for the early diagnosis of HCC patients, corresponding to 10.0% for AFP-L3 (26). Therefore, a more definitive assessment should be performed in a larger sample set.

The MRM-MS assay and LiBA showed good agreement with regard to AFP concentration (R = 0.895), and that for AFP-L3 concentrations was moderate (R = 0.627). The discrepancy between measurements by LiBA and MRM-MS assay might have originated from matrix effects and antibody or lectin cross-reactivity. The difference in the number of calibrators (8 versus 2 points) and standard differences (internal versus external) might also have affected the accuracy of the assay. However, the reasons for the disparate performance in AFP-L3 between the 2 assays are unknown. Apparently, the imprecision in the measurements was greater at lower concentrations (eg, MRM-MS). Thus, the higher proportion of low AFP-L3 concentrations.

There are 3 notable aspects of the study that distinguish it from earlier efforts to quantify (glyco)protein biomarkers by mass spectrometry. This assay uses a monoclonal antibody to selectively enrich AFP from serum, improving the analytical sensitivity by removing most of the matrix proteins that cause interference. In general, commercially available protein antibodies are cheaper than peptide

antibodies (e.g., stable isotope standards and capture by anti-peptide antibodies) *(13)*, rendering them more cost-effective in quantifying low-concentration (glyco)proteins in blood.

One of the challenges of bottom-up proteomics is that quantifying endogenous proteins in human serum depends highly on the sample preparation, which is subject to extreme variability *(3, 14)*. To minimize this variability, I used a stable isotope-labeled internal standard protein analog. The advantage of using such compounds instead of peptide analogs is that they can be added at the beginning of the sample preparation, mitigating variations in enrichment, fractionation, deglycosylation, and digestion *(15)*. In addition, all sample preparation steps were performed on a volumetric basis ("addition only") to enhance their compatibility with liquid handling systems.

I validated this analytical method in accordance with integrated multinational guidelines. This study is the first to satisfy the experimental design requirements set by the guidelines of major institutes. Table 2-17 details the criteria that dictated the experimental design with respect to the validation of the analytical method. In addition, the schedule constitutes an efficient program for conducting MRM-MS analyses, minimizing the number of injections into the equipment. I hope that this proof-of-principle application of the suggested guidelines will be helpful for validating other biomarkers as they progress through the clinical application pipeline.

In summary, the method quantifies AFP-L3, a biomarker of HCC, with greater sensitivity than LiBA. I recommend implementation of this MRM-MS

assay, which is superior in distinguishing HCC from non-HCC versus LiBA, despite it being a fundamentally disparate method from the conventional AFP-L3 assay. The strategies that I have used are highly transferable to other studies of (glyco)protein biomarkers that have significant clinical value but are difficult to quantify due to the low sensitivity of conventional equipment.

V. CONCLUSIONS

In this study, I have developed and validated an MRM-MS assay for quantifying AFP, AFP-L3%, and PIVKA-II in human serum to diagnose early-stage HCC. Especially, the improved analytical sensitivity of this MRM-MS assay allows us to detect AFP-L3 concentrations that are not observable by LiBA. Also, I identified the surrogate peptide to measure PIVKA-II in human serum by mass spectrometry. This study improved the low analytical sensitivity and low reproducibility of previous study (16) by immunoprecipitation of AFP using a monoclonal antibody and separately measuring AFP and AFP-L3 after fractionation with LCA lectin. This process resulted in a wider quantification range than all other AFP (0.051-4000 ng/mL), AFP-L3 (0.132%-100%) and PIVKA-II (1.28-100000 ng/mL) measurement method. With regard to its application in clinical practice, this method validation study, performed in accordance with several international guidelines, such as those from the FDA, EMA, KFDA, and CLSI, demonstrated that the assay is robust and reproducible. This SRM-MS assay is immediately applicable to clinical practice. Future interlaboratory work will be performed to confirm that this method is robust and reproducible across laboratories.

VI. REFERENCES

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.
2. Gomaa AI, Khan SA, Toledano MB, Waked I, Taylor-Robinson SD. Hepatocellular carcinoma: Epidemiology, risk factors and pathogenesis. World journal of gastroenterology : WJG 2008;14:4300-8.
3. Xiao JF, Varghese RS, Zhou B, Nezami Ranjbar MR, Zhao Y, Tsai TH, et al. Lc-ms based serum metabolomics for identification of hepatocellular carcinoma biomarkers in egyptian cohort. Journal of proteome research 2012;11:5914-23.
4. Shah DV, Swanson JC, Suttie JW. Abnormal prothrombin in the vitamin k-deficient rat. Thrombosis research 1984;35:451-8.
5. Beale G, Chattopadhyay D, Gray J, Stewart S, Hudson M, Day C, et al. Afp, pivkaii, gp3, scca-1 and follisatin as surveillance biomarkers for hepatocellular cancer in non-alcoholic and alcoholic fatty liver disease. BMC cancer 2008;8:200.
6. Kasahara A, Hayashi N, Fusamoto H, Kawada Y, Imai Y, Yamamoto H, et al. Clinical evaluation of plasma des-gamma-carboxy prothrombin as a marker protein of hepatocellular carcinoma in patients with tumors of various sizes. Dig Dis Sci 1993;38:2170-6.
7. Shimada M, Yamashita Y, Hamatsu T, Hasegawa H, Utsunomiya T, Aishima S, Sugimachi K. The role of des-gamma-carboxy prothrombin levels in hepatocellular carcinoma and liver tissues. Cancer letters 2000;159:87-94.
8. Sakamoto N. Nx-pvka assay, a conventional but refined prognostic biomarker for hepatocellular carcinoma. Journal of gastroenterology and hepatology 2013;28:755-6.
9. Pote N, Cauchy F, Albuquerque M, Voitot H, Belghiti J, Castera L, et al. Performance of pivka-ii for early hepatocellular carcinoma diagnosis and prediction of microvascular invasion. J Hepatol 2015;62:848-54.

10. Toyoda H, Kumada T, Osaki Y, Tada T, Kaneoka Y, Maeda A. Novel method to measure serum levels of des-gamma-carboxy prothrombin for hepatocellular carcinoma in patients taking warfarin: A preliminary report. Cancer science 2012;103:921-5.
11. Friedman PA, Przysiecki CT. Vitamin k-dependent carboxylation. Int J Biochem 1987;19:1-7.
12. Huisse MG, Leclercq M, Belghiti J, Flejou JF, Suttie JW, Bezeaud A, et al. Mechanism of the abnormal vitamin k-dependent gamma-carboxylation process in human hepatocellular carcinomas. Cancer 1994;74:1533-41.
13. Blanchard RA, Furie BC, Jorgensen M, Kruger SF, Furie B. Acquired vitamin k-dependent carboxylation deficiency in liver disease. N Engl J Med 1981;305:242-8.
14. Motohara K, Endo F, Matsuda I. Effect of vitamin k administration on acarboxy prothrombin (pivka-ii) levels in newborns. Lancet 1985;2:242-4.
15. Motohara K, Kuroki Y, Kan H, Endo F, Matsuda I. Detection of vitamin k deficiency by use of an enzyme-linked immunosorbent assay for circulating abnormal prothrombin. Pediatr Res 1985;19:354-7.
16. Uehara S, Gotoh K, Handa H, Tomita H, Senshuu M. Distribution of the heterogeneity of des-gamma-carboxyprothrombin in patients with hepatocellular carcinoma. Journal of gastroenterology and hepatology 2005;20:1545-52.
17. Tameda M, Shiraki K, Sugimoto K, Ogura S, Inagaki Y, Yamamoto N, et al. Des-gamma-carboxy prothrombin ratio measured by p-11 and p-16 antibodies is a novel biomarker for hepatocellular carcinoma. Cancer science 2013;104:725-31.
18. Suzuki K. Positioning of novel tumor marker nx-pvka-r in the diagnosis of hepatocellular carcinoma in comparison with pivka-ii. Dokkyo Journal of Medical Sciences 2013.
19. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions. Nat Methods 2012;9:555-

66.

20. Kettenbach AN, Rush J, Gerber SA. Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. Nature protocols 2011;6:175-86.
21. Makawita S, Diamandis EP. The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: Current strategies for candidate verification. Clinical chemistry 2010;56:212-22.
22. Kinukawa H, Shirakawa T, Yoshimura T. Epitope characterization of an anti-pivka-ii antibody and evaluation of a fully automated chemiluminescent immunoassay for pivka-ii. Clinical biochemistry 2015.
23. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: The long and uncertain path to clinical utility. Nat Biotechnol 2006;24:971-83.
24. Taketa K, Sekiya C, Namiki M, Akamatsu K, Ohta Y, Endo Y, Kosaka K. Lectin-reactive profiles of alpha-fetoprotein characterizing hepatocellular carcinoma and related conditions. Gastroenterology 1990;99:508-18.
25. Zhu K, Dai Z, Zhou J. Biomarkers for hepatocellular carcinoma: Progression in early diagnosis, prognosis, and personalized therapy. Biomark Res 2013;1:10.
26. Campbell MP, Peterson R, Mariethoz J, Gasteiger E, Akune Y, Aoki-Kinoshita KF, et al. Unicarbkb: Building a knowledge platform for glycoproteomics. Nucleic acids research 2014;42:D215-21.
27. Leerapun A, Suravarapu SV, Bida JP, Clark RJ, Sanders EL, Mettler TA, et al. The utility of lens culinaris agglutinin-reactive alpha-fetoprotein in the diagnosis of hepatocellular carcinoma: Evaluation in a united states referral population. Clin Gastroenterol Hepatol 2007;5:394-402; quiz 267.
28. Wang SS, Lu RH, Lee FY, Chao Y, Huang YS, Chen CC, Lee SD. Utility of lentil lectin affinity of alpha-fetoprotein in the diagnosis of hepatocellular carcinoma. J Hepatol 1996;25:166-71.

29. Kumada T, Nakano S, Takeda I, Kiriyama S, Sone Y, Hayashi K, et al. Clinical utility of lens culinaris agglutinin-reactive alpha-fetoprotein in small hepatocellular carcinoma: Special reference to imaging diagnosis. J Hepatol 1999;30:125-30.
30. Simultaneous determination of percentage of lens culinaris agglutinin-reactive a-fetoprotein and a-fetoprotein concentration using the libasys clinical auto-analyzer. Clinica Chimica Acta 2003.
31. Fuzery AK, Levin J, Chan MM, Chan DW. Translation of proteomic biomarkers into fda approved cancer diagnostics: Issues and challenges. Clin Proteomics 2013;10:13.
32. Kurosawa T, Watanabe M. Development of on-chip fully automated immunoassay system "mutaswako i30" to measure the changes in glycosylation profiles of alpha-fetoprotein in patients with hepatocellular carcinoma. Proteomics 2016;16:3056-61.
33. Tamura Y, Igarashi M, Kawai H, Suda T, Satomura S, Aoyagi Y. Clinical advantage of highly sensitive on-chip immunoassay for fucosylated fraction of alpha-fetoprotein in patients with hepatocellular carcinoma. Dig Dis Sci 2010;55:3576-83.
34. Toyoda H, Kumada T, Tada T, Kaneoka Y, Maeda A, Kanke F, Satomura S. Clinical utility of highly sensitive lens culinaris agglutinin-reactive alpha-fetoprotein in hepatocellular carcinoma patients with alpha-fetoprotein <20 ng/ml. Cancer Sci 2011;102:1025-31.
35. Hoofnagle AN, Becker JO, Oda MN, Cavigiolio G, Mayer P, Vaisar T. Multiple-reaction monitoring-mass spectrometric assays can accurately measure the relative protein abundance in complex mixtures. Clin Chem 2012;58:777-81.
36. Ackermann BL, Berna MJ. Coupling immunoaffinity techniques with ms for quantitative analysis of low-abundance protein biomarkers. Expert Rev Proteomics 2007;4:175-86.
37. Dufield DR, Radabaugh MR. Online immunoaffinity lc/ms/ms. A general

method to increase sensitivity and specificity: How do you do it and what do you need? Methods 2012;56:236-45.

38. Kushnir MM, Rockwood AL, Roberts WL, Abraham D, Hoofnagle AN, Meikle AW. Measurement of thyroglobulin by liquid chromatography-tandem mass spectrometry in serum and plasma in the presence of antithyroglobulin autoantibodies. Clin Chem 2013;59:982-90.
39. Liebman HA, Furie BC, Tong MJ, Blanchard RA, Lo KJ, Lee SD, et al. Des-gamma-carboxy (abnormal) prothrombin as a serum marker of primary hepatocellular carcinoma. N Engl J Med 1984;310:1427-31.
40. Fujiyama S, Morishita T, Sagara K, Sato T, Motohara K, Matsuda I. Clinical evaluation of plasma abnormal prothrombin (pivka-ii) in patients with hepatocellular carcinoma. Hepatogastroenterology 1986;33:201-5.
41. Revermann T, Gotz S, Kunnemeyer J, Karst U. Quantitative analysis by microchip capillary electrophoresis: Current limitations and problem-solving strategies. Analyst 2008;133:167-74.
42. Altria KD, Bestford J. Main component assay of pharmaceuticals by capillary electrophoresis: Considerations regarding precision, accuracy, and linearity data. J Capillary Electrophor 1996;3:13-23.
43. Wielgos T, Turner P, Havel K. Validation of analytical capillary electrophoresis methods for use in a regulated environment. J Capillary Electrophor 1997;4:273-8.
44. Hoofnagle AN, Becker JO, Wener MH, Heinecke JW. Quantification of thyroglobulin, a low-abundance serum protein, by immunoaffinity peptide enrichment and tandem mass spectrometry. Clin Chem 2008;54:1796-804.
45. Kim H, Kim K, Yu SJ, Jang ES, Yu J, Cho G, et al. Development of biomarkers for screening hepatocellular carcinoma using global data mining and multiple reaction monitoring. PLoS One 2013;8:e63468.
46. Kim H, Kim K, Jin J, Park J, Yu SJ, Yoon JH, Kim Y. Measurement of glycosylated alpha-fetoprotein improves diagnostic power over the native form in hepatocellular carcinoma. PLoS One 2014;9:e110366.

47. Kim H, Park J, Kim Y, Sohn A, Yeo I, Jong Yu S, et al. Serum fibronectin distinguishes the early stages of hepatocellular carcinoma. Sci Rep 2017;7:9449.
48. Kim H, Yu SJ, Yeo I, Cho YY, Lee DH, Cho Y, et al. Prediction of response to sorafenib in hepatocellular carcinoma: A putative marker panel by multiple reaction monitoring-mass spectrometry (mrm-ms). Mol Cell Proteomics 2017;16:1312-23.
49. Yu SJ, Kim H, Min H, Sohn A, Cho YY, Yoo JJ, et al. Targeted proteomics predicts a sustained complete-response after transarterial chemoembolization and clinical outcomes in patients with hepatocellular carcinoma: A prospective cohort study. J Proteome Res 2017;16:1239-48.
50. Huttenhain R, Soste M, Selevsek N, Rost H, Sethi A, Carapito C, et al. Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. Sci Transl Med 2012;4:142ra94.
51. Shi T, Su D, Liu T, Tang K, Camp DG, 2nd, Qian WJ, Smith RD. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics 2012;12:1074-92.
52. Sohn A, Kim H, Yu SJ, Yoon JH, Kim Y. A quantitative analytical method for pivka-ii using multiple reaction monitoring-mass spectrometry for early diagnosis of hepatocellular carcinoma. Anal Bioanal Chem 2017;409:2829-38.
53. Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, et al. Isotope-labeled protein standards: Toward absolute quantitative proteomics. Mol Cell Proteomics 2007;6:2139-49.
54. Ahn DG, Kim HJ, Kang H, Lee HW, Bae SH, Lee JH, et al. Feasibility of alpha-fetoprotein as a diagnostic tool for hepatocellular carcinoma in korea. The Korean journal of internal medicine 2016;31:46-53.
55. Korean Liver Cancer Study G, National Cancer Center K. [practice guidelines for management of hepatocellular carcinoma 2009]. The Korean journal of hepatology 2009;15:391-423.
56. Naraki T, Kohno N, Saito H, Fujimoto Y, Ohhira M, Morita T, Kohgo Y.

Gamma-carboxyglutamic acid content of hepatocellular carcinoma-associated des-gamma-carboxy prothrombin. Biochim Biophys Acta 2002;1586:287-98.

57. Hoofnagle AN, Whiteaker JR, Carr SA, Kuhn E, Liu T, Massoni SA, et al. Recommendations for the generation, quantification, storage, and handling of peptides used for mass spectrometry-based assays. Clin Chem 2016;62:48-69.
58. Uehara S, Gotoh K, Handa H, Honjo K, Hirayama A. Process of carboxylation of glutamic acid residues in the gla domain of human des-gamma-carboxyprothrombin. Clin Chim Acta 1999;289:33-44.
59. Yu R, Xiang X, Tan Z, Zhou Y, Wang H, Deng G. Efficacy of pivka-ii in prediction and early detection of hepatocellular carcinoma: A nested case-control study in chinese patients. Sci Rep 2016;6:35050.
60. Lou J, Zhang L, Lv S, Zhang C, Jiang S. Biomarkers for hepatocellular carcinoma. Biomark Cancer 2017;9:1-9.
61. Ferrin G, Aguilar-Melero P, Rodriguez-Peralvarez M, Montero-Alvarez JL, de la Mata M. Biomarkers for hepatocellular carcinoma: Diagnostic and therapeutic utility. Hepat Med 2015;7:1-10.

ABSTRACT IN KOREAN

국문초록

서론: 혈액 내의 AFP, AFP-L3 그리고 PIVKA-II 는 간세포암의 표지자이다. 위의 3 가지 단백질 표지자는 liquid-phase binding assay (LiBA)를 이용해서 정량 하는 것이 일반적이다. 그러나 LiBA 는 기계 자체의 분석적인 민감도가 낮기 때문에 표지자의 농도를 정확하게 반영할 수 없다는 단점이 있다. 그러므로 우리는 질량분석기를 이용한 다중 반응 검지법을 이용하여 AFP, AFP-L3 그리고 PIVKA-II 를 정량 할 수 있는 정량법을 개발 하는 것을 목표로 연구하였다.

방법: 안정 동위 원소로 표지된 재조합 단백질을 내부 표준 물질로 이용하였고, 모노클론항체를 이용하여 AFP 와 prothrombin 을 인리치 하였고, LCA 렉틴을 이용하여 AFP-L3 만을 분획하였으며 이 후, 트립신 (AFP, AFP-L3)과 키모트립신 (PIVKA-II)으로 다이제스천을 시행한 후 질량분석기에 주입하여 online-desalting, MRM-MS 분석을 진행하였다. 1 장에서, 400 명의 시료 (간염 100, 간경화 100, 간암 200)를 LiBA 와 MRM-MS 로 각각 분석한 후, 두 정량법의 진단적 성능을 비교하였다. 2 장에서, 질량분석기 기반으로 개발된 정량법을 국제적인 가이드라인 (FDA, EMA, KFDA, CLSI)을 따라서 분석적 성능을 검증하였다.

결과: 1 장에서, 개발한 MRM-MS 기반의 정량법의 AFP 와 AFP-L3 최저정량한계는 0.051 ng/mL 로써 LiBA 의 최저정량한계보다 낮았다. 따라서, LiBA 로는 검출하지 못했던 39 명의 간암 환자를 MRM-MS 정량법을 통해서는 검출할 수 있었다. PIVKA-II 의 경우, MRM-MS 로 PIVKA-II 의 농도를 정량하기 위한 최적의 타겟 펩타이드를 선별하였다. 그리고 1.28 - 100000 ng/mL 의 농도 구간에서 선형성을 확보하였다. 2 장에서는 개발된 MRM-MS 정량법을 12 가지 항목 (검량선, 분석적 특이도, 분석적 민감도, 캐리오버, 정밀도, 분석법의 회수율, 생체시료회수율, 면역침강반응의 회수율, 희석타당도, 안정성, 재현성, 품질관리시료)을 국제적인 가이드라인에 분석적 검증을 하였고, 대부분의 항목이 가이드라인의 기준에 부합하였다.

결론: 우리는 혈액 속의 AFP, AFP-L3 그리고 PIVKA-II 를 정량하기 위해서 액체 크로마토그래피와 다중 반응 검지법 기반의 분석법을 개발하였고, 이 분석법은 기존의 LiBA 분석법의 낮은 민감도를 극복할 수 있다. AFP-L3 의 분석적 민감도를 향상시키기 위해서 우리는 단일클론항체와 렉틴을 이용하여 AFP-L3 를 분획하였고, 이 후에 탈당화와 펩타이드화를 진행하였다. PIVKA-II 의 경우, 혈액속의 PIVKA-II 를 질량분석기로 정량 하기 위한 타겟 펩타이드를 선별하여, MRM-MS 기반의 분석법을 개발하였다. 개발된 분석법은 국제적인 가이드라인 (FDA, EMA, KFDA, CLSI)을 따라서 분석적 검증을 완료하여 충분히 정확하고 재현성있게 정량할 수 있는 분석법인 것을 확인하였다.

Keywords: 다중반응검지법; 바이오마커; 알파태아단백질; 당질화된 알파태아단백질; 피브카; 분석적 검증; 간세포암;

Student number: 2014-21997

* 본 내용은 Analytical and Bioanalytical Chemistry 학술지에 출판 완료된 내용임. A quantitative analytical method for PIVKA-II using multiple reaction monitoring-mass spectrometry for early diagnosis of hepatocellular carcinoma. Analytical and Bioanalytical Chemistry 10.1007/s00216-017-0227-8.

* 본 내용은 Journal of Pharmaceutical and Biomedical Analysis 학술지에 출판 완료된 내용임. Fully validated SRM-MS-based method for absolute quantification of PIVKA-II in human serum: Clinical applications for patients with HCC. Journal of Pharmaceutical and Biomedical Analysis Journal. 10.1016/j.jpba.2018.04.025.

* 본 내용은 Clinical Chemistry 학술지에 출판 완료된 내용임. Clinical Assay for AFP-L3 by Using Multiple Reaction Monitoring–Mass Spectrometry for Diagnosing Hepatocellular Carcinoma Clinical chemistry. 1373/clinchem.2018.289702.